A Quarterly Journal Devoted to All Aspects of Cryptology

CRYPTOLOGIA

Volume 8 Number 2 April 1984

24	33 ALLIED	43	L ALLOWANCE(S)	Q ANSWER(ED)	A ARMORED	75
O BATTALION(S)	A	E BEARING	50	R BRIDGE	66 BRIGADE	G BRITISH
V	SWITCH ON	44 CENTER	51 CHANGE	60 CHART	B	76 CIRCUIT
25 CONFIRM(ED)	34 CONNECT (ION)	F CONSTRUCTION	52 CONTACT(ED)	S COORDINATE(S)	SWITCH ON	H CORPS
26	35 DIRECT	G DISTRIBUTE(D)	M DIVISION(S)	61	67 EAST	77 ECHELON
27 ESTABLISH (ING) (ED)	36 ESTIMATE(D)	45 ETOUSA	53	T EXCHANGE	68 EXPECT(ED)	SWITCH OFF
W FORCE	B FORWARD	46	SWITCH ON	U FRENCH	69 FREQUENCY	I
X HELP	37	H HIGH	54 HILL	SWITCH ON	C HOLD/HELD	78 HOUR
SWITCH ON	38 INFANTRY	47 INFORM (ATION)	N	62	7 JOIN(ED)	79
28 LINE	39 LINK	L LOCATE (D) (ION)	55	V MAP	D MEET	8
29 NEAR	C	SWITCH ON	56 NORTH	W NOTHING TO REPORT	70	J OBJECT(IVE)
Y PERSONNEL	D PIGEON	48 PHOTO	57 PLATOON	63 POINT	71 POSITION	80 POST
3 RECEIVE(R)	4	J RELAY(ED)	58	X REPAIR(ED)	E	81 REPEAT(ER)
30 REUR MSG. NO.	40 RIGHT	49 ROAD	O ROW REGISTER	64 SAME	72 SECTION	K SECTOR
Z SIGN	41 SIGNAL	K SOUTH	59	Y STATION	73 STOP	L
31 TANK	E TELEGRAPH		6 TELEPHONE	65 TELETYPE	F	82
32 TRANSPORT	42	5 UNTIL	P USE	Z WEST	74	83 YESTERDAY

SIG NO. 1

CRYPTOLOGIA

A Quarterly Journal Devoted to All Aspects of Cryptology

All correspondence concerning subscriptions, advertising and publications should be sent to the publisher at the Editorial Office, Rose-Hulman Institute of Technology, Terre Haute, Indiana 47803. TELEX 752486.

See inside back cover for subscription information.

ISSN 0161 - 1194. Manufactured in the United States of America.

Cover: For more of and about this encryption device see page 163.

Assistance of Rose-Hulman Institute of Technology is acknowledged and appreciated.

CRYPTANALYSIS OF A MACLAREN-MARSAGLIA SYSTEM

CHARLES T. RETTER

ABSTRACT: A successful attack on a computer file encryption system is described. The system was based on the MacLaren-Marsaglia algorithm ([1], [2] p. 30) for generating a sequence of pseudo-random numbers from two other sequences.

INTRODUCTION

The research and development groups at Data General do much of their work on a large network of interconnected minicomputers. In this environment (see [3]), it is not practical to rely on the operating system to protect files from unauthorized access, simply because most people have physical access to the unattended machines. For this reason, various file encryption systems have been developed. The early versions were trivial, but by 1980 a program was in use which its author claimed to be "virtually unbreakable short of exhaustive search." Since the key size was 31 bits, exhaustive search might have been possible, but on the available minicomputers it would have taken days of CPU time even with known plaintext. The system proved to be far less secure, and can usually be broken in minutes using just a guess about the plaintext.

THE ENCRYPTION ALGORITHM

A simple disassembly of the program revealed the algorithm (and in fact produced a better understanding of what it really was doing than reading the source would have, since there were several mistakes in the source). The following is a somewhat simplified description of the algorithm. First two linear congruential generators are defined:

```
FUNCTION RANDOM1;
BEGIN
   SEED1 := (46876 * SEED1 + 32749) MOD 59049;
   RETURN(SEED1);
END;
```

```
FUNCTION RANDOM 2;
BEGIN
   SEED2  := ( 4353 * SEED2 + 32633) MOD 32768;
RETURN (SEED2);
END;
```

Now a third function is defined, using a table of 257 elements.

```
FUNCTION RANDOM3;
BEGIN
   I := RANDOM1 MOD 257
   SEED3 := TABLE[I]
   TABLE [I] := RANDOM2
   RETURN(SEED3);
   END;
```

This algorithm is a version of the MacLaren-Marsaglia algorithm, which Knuth [2] contends "will satisfy virtually anyone's requirements for randomness". The algorithm is used by first initializing SEED1 and SEED2 with values that serve as the key. Next, the function RANDOM3 is called some number of times, which happens to depend on the file length. Then each 16-bit word in the file is exclusive-ORed with a word produced by RANDOM3. Obviously, the same procedure can be used either to encipher or to decipher a file.

The periods of the sequences produced by RANDOM1 and RANDOM2 are easily determined. Both functions satisfy the conditions to generate maximal length sequences ([2] p. 15), so the period of RANDOM1 is 59049 and the period of RANDOM2 is 32768. Since these two numbers are relatively prime, the period of RANDOM3 will be 1934917632.

BREAKING THE ALGORITHM

The method of attack used was the known-plaintext attack. Since it turns out that about ten known characters are usually sufficient, it is possible to use a guess as the known plaintext. Assuming that the plaintext is correct, a sequence of words from RANDOM3 becomes available. If we can determine the values from RANDOM1 and RANDOM2 which produced these words, we will be able to reconstruct the state of the table, and calculate the initial values of the seeds.

Notice that all of the numbers produced by the function RANDOM3 actually were generated by RANDOM2. The only effect of RANDOM1 is to insert a varying delay between the time that RANDOM2 generates a number and the time that RANDOM3

uses it to transform the file. Using the method described in the appendix, we can quickly calculate the time that a given number was generated, relative to some arbitrary starting point. Of course, we don't know the actual starting point (that was part of the key), but the differences between these times of generation will be independent of the starting point. In fact, the sequence of differences between generation times (from RANDOM2) of the numbers produced by RANDOM3 depends only on the state of RANDOM1, not on RANDOM2.

Since the period of RANDOM1 is 59049, the period of this sequence of differences will also be 59049, and the sequence can quickly be generated and placed in a table. The table will be accessed by specifying a short subsequence. A subsequence of four differences will uniquely identify a point in the overall sequence in 97% of all cases. A subsequence of 11 differences is required for uniqueness in all cases. When the table is generated, the values of RANDOM1 and the delays through the table can also be generated and stored.

The general procedure for breaking the algorithm can now be described. Let the values produced by the generators at time i be denoted RANDOM1[i], RANDOM2[i], and RANDOM3[i]. Then proceed as follows:

(1) Take the string assumed to be plaintext and XOR it with the ciphertext at some position in the file. If this is the correct position, the results will be RANDOM3[i], RANDOM3[i], ...

(2) Since each of the RANDOM3 values was generated by RANDOM2 at some time, there exist values of j and $\Delta[i]$ such that

$$\text{RANDOM3}[i] = \text{RANDOM2}[j]$$

$$\text{RANDOM3}[i+1] = \text{RANDOM2}[j+\Delta[i]]$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

where $\Delta[i]$ is a function of RANDOM1[i].

(3) Using the procedure in Appendix B, calculate the values j, $j+\Delta[i]$, $j+\Delta[i]+\Delta[i+1]$, $j+\Delta[i]+\Delta[i+1]+\Delta[i+2]$, and $j+\Delta[i]+\Delta[i+1]+\Delta[i+2]+\Delta[i+3]$. Then the consecutive differences between these values produce $\Delta[i]$, $\Delta[i+1]$, $\Delta[i+2]$, and $\Delta[i+3]$. If the magnitude of any of these differences is greater than 2787, go back to step (1) and try another position for the guess, since 2787 is the largest possible Δ.

(4) Search the Δ sequence (which may be generated once using RANDOM1 and the TABLE and placed in a hash table) for the four consecutive values just

found. If the values are not found, go back to step (1) and try another position for the guess. If the values are found, and are unique, the value of RANDOM1[i] is known. Occasionally, the four values will not be unique. In that case, either additional values must be used or each of the possible matches must be tried, and the resulting plaintext checked.

(5) Since RANDOM1 also determines how long each element of RANDOM2 spends in the TABLE, RANDOM1[i] can be used to find i-j, i.e. the length of time that RANDOM2[j] spent in the TABLE. Since i is known, from the position in the file and the initialization delay, j can now be calculated.

(6) Use i and RANDOM1[i] to calculate the initial value of SEED1.

(7) Use j and RANDOM2[j] to calculate the initial value of SEED2.

Because the calculation of the differences is very fast, and the differences eliminate almost all incorrect guesses, it isn't necessary to improve the procedure. Most files are broken in a minute or two assuming that the guess is found within the first thousand character positions. See Appendix C for an example of this procedure.

DOUBLE-ENCRYPTION

When the author of the encryption program was informed that it had been broken, he made several attempts to improve it. In general, these attempts did not make the cipher much more difficult to break. It should be clear from the above procedure that even a significant increase in the key size would not make the system secure. Its primary weakness is that the effects of the two generators RANDOM1 and RANDOM2 can be separated and attacked individually. Furthermore, most guesses can be eliminated by using only the inverse of RANDOM2, which can be computed efficiently as shown in the appendix.

The most successful countermeasure seemed to be double-encryption. Double-encryption strengthens many systems, (see [4], for example), and in this case results in an effective key size of 62 bits, ruling out a brute-force attack. Also, double-encryption seems to remove the weakness that was used in the above attack, since with double-encryption the sequence of numbers obtained by exclusive-ORing the known plaintext with the ciphertext no longer consists of numbers from the RANDOM2 sequence. Instead, each value is the XOR of two numbers from differently delayed RANDOM2 sequences.

Notice, however, that the modulus of RANDOM2 is a power of 2. This implies that the least significant bits of the numbers in the RANDOM2 sequence must

either be constant or alternating 0s and 1s ([2] p. 12). In fact, they are alternating 0s and 1s. Therefore, the XOR of any two numbers from the RANDOM2 sequence will be even if the difference in their times of generation was even, and odd if the difference in their times of generation was odd. Since the delay through the table is a function only of RANDOM1, the least-significant bits of successive XORs of the values produced by RANDOM3 form a pattern which identifies the state of RANDOM1, a cryptographic weakness, as shown below.

The period of this sequence of bits is 59049, so the sequence can easily be generated and stored in a table. However, considerably more known-plaintext will be required to ensure that the XOR of two subsequences of this sequence is unique. Using subsequences of 38 bits results in a probability of uniqueness greater than 99%. If enough plaintext is available, the use of 64-bit subsequences will guarantee uniqueness, but duplicate matches are not a serious problem since they can be detected and the correct choice made based on the values of the RANDOM2 sequence.

The procedure for breaking a double-encrypted message is as follows:

(1) Take 78 characters of known plaintext, and XOR them with the ciphertext to produce a sequence of 39 numbers, which should be equal to the XORs of two RANDOM3 sequences. Let the least significant bits of the known sequence be called R[i], R[i+1], etc., and let the least significant bits of the two unknown RANDOM3 sequences be called

$$R3[i],\ R3[i+1],\ldots \quad \text{and} \quad r3[i],\ r3[i+1],\ldots$$

(2) Then we know that

$$R[i] = R3[i] \oplus r3[i] = R2[j] \oplus r2[k]$$

$$R[i+1] = R3[i+1] \oplus r3[i+1] = R2[j+\Delta[i]] \oplus r2[k+\delta[i]]$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

So, if we XOR successive values of the R sequence, we get

$$R[i] \oplus R[i+1] = R2[j] \oplus R2[j+\Delta[i]] \oplus r2[k] \oplus r2[k+\delta[i]]$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

Since $R2[x] \oplus R2[x+y] = y \bmod 2$ for any x and y, we have

$$R[i] \oplus R[i+1] = \Delta[i] \bmod 2 \oplus \delta[i] \bmod 2$$
$$R[i+1] \oplus R[i+2] = \Delta[i+1] \bmod 2 \oplus \delta[i+1] \bmod 2$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

(3) Using a table of sequences of (Δ mod 2) which are produced by the RANDOM1 generator, pick one 38-bit sequence and XOR it with the 38-bit sequence found in the previous step. If the resulting sequence is also in the table, proceed to the next step. Otherwise, try the next 38-bit sequence until a match is found. There are a total of 59049 entries in the table, since that is the period of RANDOM1. Naturally, the table should be organized to make this step fast.

(4) When the two sequences have been found in the table, the values of RANDOM1[i], random1[i], and (i-j) and (i-k) can also be found, since they depend only on the RANDOM1 generator. These values should also have been placed in the table when it was generated.

(5) Assume a value for RANDOM2[j]. Using this value, and the values of Δ[i], Δ[i+1], etc., calculate the succeeding values RANDOM2[j+Δ[i]], RANDOM2[j+Δ[i]+Δ[i+1]], etc. Then use these values and the original values obtained from the plaintext-ciphertext to calculate random2[k], random2[k+δ[i]], random2[k+δ[i]+δ[i+1]], etc. Compare these values against the known sequence δ[i], δ[i+1], etc. If they match, the values of RANDOM2[j] and random2[k] have been found. If not, try another guess for RANDOM2[j]. There are 32768 possible values for RANDOM2[j], so this step should be done as efficiently as possible. Almost all of the incorrect values of RANDOM2[j] can be eliminated by calculating only a single value of δ. Note that the δ sequence can also be used to distinguish between duplicate matches in step (3).

(6) Using RANDOM1[i] and i calculate the initial value of SEED1.

(7) Using random1[i] and i calculate the initial value of seed1.

(8) Using RANDOM2[j] and j calculate the initial value of SEED2.

(9) Using random2[k] and k calculate the initial value of seed2.

In summary, double-encryption makes the system more difficult to attack, since more known-plaintext is required, larger tables are used, and each step takes about 50,000 times as long. However, the system is still fairly easy to break, even though the key size is now greater than the 56-bit key used for DES.

CONCLUSION

After a number of attempts to improve the algorithm, the author of the file encryption program finally gave up and replaced the whole program with a similar one based on DES. Although a software implementation has obvious problems, the DES program is not likely to be broken as easily as this MacLaren-Marsaglia program was.

The primary weakness of the MacLaren-Marsaglia algorithm as a cipher system is that the effects of the two constituent generators can be separated. This has the effect of halving the key size. The attack described in this paper used certain features which actually improve the statistics of the pseudo-random numbers. For example, the fact that RANDOM1 has a potency of 10 improves its randomness, but it also reduces the number of differences required to uniquely identify its state. Similarly, RANDOM2 has a very high potency, but its least-significant bits are not at all random. However, changing the generators to remove these weaknesses would probably result in the introduction of new weaknesses which could be exploited in similar ways.

It may be worth noting that a recent article [5] describes a commercial cryptographic device as using the MacLaren-Marsaglia algorithm. The experience described above leads me to question the security of such a device.

REFERENCES

1. MacLaren, M.D. and G. Marsaglia, 1965. Uniform Random Number Generators. J. ACM 12: 83-89.

2. Knuth, D.E. 1969. The Art of Computer Programming. Volume 2 - Seminumerical Algorithms. Reading, MA: Addison-Wesley.

3. Kidder, T. 1981. The Soul of a New Machine. Boston: Little, Brown.

4. Merkle, R.C. and M.E. Hellman. 1981. On the Security of Multiple Encryption. CACM 24: 465-467.

5. Kruh, L. 1983. Cipher Equipment: The Cryptographic Unit CSI-10. Cryptologia. 7: 83-88.

Appendix A

Efficient Computation of Values in a Linear Congruential Sequence

Although it is possible to generate a given value in a linear congruential sequence simply by generating each successive value until the desired one is obtained, this method requires a significant amount of time. Another possibility is to generate all of the values once, place them in a table, and extract the desired one when it is needed. This method requires less time, but may take a large amount of space in the computer's memory. Fortunately, by taking advantage of the linearity of the sequence, values may be obtained using small tables and a small amount of computation.

Let the sequence be defined by the following recursion, with $X_0 = 0$:

$$X_{n+1} = (aX_n + b) \bmod m .$$

Then
$$X_n = \sum_{i=0}^{n-1} a^i b \bmod m,$$

and
$$X_{n+j} - X_n = \sum_{i=n}^{n+j-1} a^i b \bmod m,$$

so
$$X_{n+j} = X_n + a^n X_j \bmod m.$$

Now, assume that we have tables of X_n and a^n for $n+1,2,4,8,16,\ldots$. Then, given any value of X_j, we can calculate the value of X_{j+1} or X_{j+2} or X_{j+4}, etc., with only one multiplication and one addition (mod m). By repeating this procedure, we can calculate the value of any X in the sequence with at most 16 multiplications, 16 additions, and possibly 16 divisions, since all of the numbers being used here are limited to 16 bits. The memory space required for the tables is negligible, since there are only 16 numbers in each table.

The following procedure can be used to find any value in the RANDOM2 sequence, given any other value in the sequence, and the difference between their times of generation. Note that no divisions are required because the modulus is a power of 2.

```
FUNCTION R2VALUE(STARTVALUE,OFFSET);
BEGIN
FOR I:=0 TO 14 DO
    IF (OFFSET AND 2↑I)
    THEN STARTVALUE := A[I] * STARTVALUE + X[I];
RETURN (STARTVALUE AND 32767);
END;
```

APPENDIX B

Efficient Computation of the Inverse of the RANDOM2 Sequence

The inverse of the previous function, namely, finding the difference between the times of generation of two known values from the RANDOM2 sequence, can be computed with a similar amount of effort. However, the method relies on certain other properties of the RANDOM2 sequence, namely, that the modulus is a power of 2, and a(mod 4) = 1. Generators with these values produce sequences with good randomness properties (see [2]), but the tables of A[I] and X[I] used in the previous section have a very convenient characteristic. Using the same notation,

$$X[I] = x^{2^i} \equiv b \sum_{j=0}^{2^i-1} a^j = \frac{a^{2^i} - 1}{a-1}$$

Using the fact that a(mod 4) = 1, and expanding, it can be shown that

$$X[I] = b\ (\ 2k + 1\ )\ 2^i \text{ for some integer } k.$$

Since b is odd, this implies that X[I] is divisible by 2^i, but not by 2^{i+1}, which means that the least significant bit which is a 1 in X[I] is the I-th bit.

Let $i = 2^I$ and $j=2^J$, where $I < J$.

Then, $$x_{i+j} \equiv X[I] + A[I]\ X[J] \bmod m.$$

Since m is a power of 2, the least significant 1 in the value x_{i+j} is the same as the least significant bit in X[I]. Therefore, given any two values from the RANDOM2 sequence, we can calculate the difference between their times of generation with the following procedure:

```
FUNCTION R2INVERSE(START,END);
BEGIN
OFFSET:=0;
FOR I:=0 TO 14 DO
    IF ( (START XOR END) AND 2↑I )
    THEN BEGIN
         OFFSET := OFFSET + 2↑I;
         START := A[I] * START + X[I];
         END;
RETURN(OFFSET);
END;
```

An assembly-language version of this routine takes about 100 microseconds on an Eclipse S/280. Since it is impossible for two consecutive values in the RANDOM3 sequence to be more than 2787 apart in the RANDOM2 sequence, most incorrect guesses can be eliminated in a fraction of a millisecond using this routine.

APPENDIX C

An Example

We are given a file of ciphertext beginning with the following values: 29400, 11661, 7238, 25666, 16219. From other information, we suspect that the first word in the file may be PROCEDURE. Then we perform the following steps to find the values of the keys:

(1) Assuming that the plaintext is correct, we can obtain five values from the RANDOM3 sequence by XORing the plaintext and the ciphertext.

```
PR -> 20562  XOR  29400 =    8842
OC -> 20291  XOR  11661 =   25294
ED -> 17732  XOR   7238 =   22786
UR -> 21842  XOR  25666 =   12560
E  -> 17696  XOR  16219 =   31355
```

(2) Since all of the values in the RANDOM3 sequence came from the RANDOM2 sequence, we know the following values of the RANDOM2 sequence:

```
RANDOM2[j]  =  8842
RANDOM2[j+Δ[i]]  =  25294
RANDOM2[j+Δ[i]+Δ[i+1]]  =  22786
RANDOM2[j+Δ[i]+Δ[i+1]+Δ[i+2]]  =  12560
RANDOM2[j+Δ[i]+Δ[i+1]+Δ[i+2]+Δ[i+3]]  =  31355
```

(3) In order to calculate the times when these values were generated by RANDOM2, we must use some arbitrary starting point. Using 0 as a starting value, and assuming that the initial value of SEED2 occurs T2 cycles after 0 in the RANDOM2 sequence, we can calculate the following values by using the function R2INVERSE defined in Appendix B.

$$j = 27994 - T2$$
$$j + \Delta[i] = 27838 - T2$$
$$j + \Delta[i] + \Delta[i+1] = 28050 - T2$$
$$j + \Delta[i] + \Delta[i+1] + \Delta[i+2] = 28048 - T2$$
$$j + \Delta[i] + \Delta[i+1] + \Delta[i+2] + \Delta[i+3] = 28051 - T2$$

Subtracting, we obtain four values from the Δ sequence:

$$\Delta[i] = -156$$
$$\Delta[i+1] = 212$$
$$\Delta[i+2] = -2$$
$$\Delta[i+3] = 3$$

Since none of these values has a magnitude greater than 2787, the guess cannot be eliminated immediately, so we proceed to the next step.

(4) In this step we search a table of the Δ sequence for the four consecutive values above. This can be done efficiently using a hash table of the subsequences, but for this example the following table shows a part of the RANDOM1 sequence and the corresponding values of Δ and the delay through TABLE (T1 is the number of cycles from 0 to the initial value of SEED1 in the RANDOM1 sequence).

t	RANDOM1[t]	mod 257	Δ[t]	table delay (i-j)
15176 - T1	43622	189	17	71
15177 - T1	49800	199	-156	55
15178 - T1	14383	248	212	212
15179 - T1	28775	248	-2	1
15180 - T1	33342	189	3	4
15181 - T1	4360	248	-83	2

Notice the Δ sequence matches, starting with the line at t = 15177 - T1.

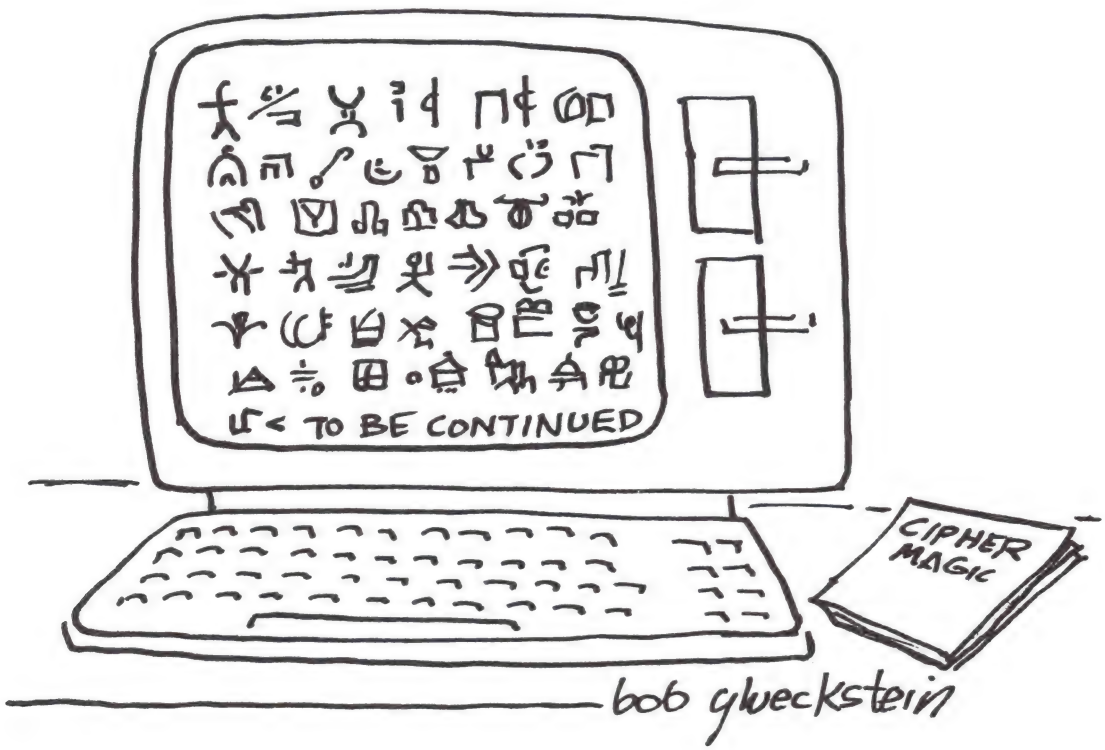
(5) From the length of the file, we know that the number of cycles of initialization was 5001. Since the ciphertext that we are using is at the beginning of the file, there is no additional offset. Therefore, the value of i is 5001. From the previous step, we know that the number

produced by RANDOM3 at that time spent 55 cycles in the TABLE. So, (i-j) = 55, and j = 5001 - 55 = 4946.

(6) We know that i = 5001 and RANDOM1[i] = 49800. To find the key, RANDOM1[0], we must go back 5001 cycles, or forward 59049 - 5001 = 54048 cycles, in the RANDOM1 sequence. Using the function explained in Appendix A, R1VALUE(49800,54048) produces the value of RANDOM1[0] = 123, which is the first key.

(7) We know that j = 4946 and RANDOM2[j] = 8842. To find the key, RANDOM2[0], we must go back 4946 cycles, or forward 32768 - 4946 = 27822 cycles, in the RANDOM2 sequence. Using the function explained in Appendix A, R2VALUE(8842,27822) produces the value of RANDOM2[0] = 456, which is the second key.

Notice that the keys could also be found by solving for the values of T1 and T2, and then finding R1VALUE(0,T1) and R2VALUE(0,T2). Knowing both keys, we can easily decipher the entire file.



PROJECT ON SECRECY AND OPENNESS IN SCIENTIFIC AND TECHNICAL COMMUNICATION

Sponsored by
American Association for the Advancement of Science
Committee on Scientific Freedom and Responsibility

In recent years, the traditional concept of scientific ideas and information as a public good, freely available to professional colleagues as well as the general public, has come under closer scrutiny. The post-World War II increase in the economic, political, and military value of scientific and technical information has fostered various private and public proposals to restrict open communication in university teaching and research activities. These proposals have cited many justifications, including national security interests, economic competition, patent protections, and quality control, as the basis for limiting access to new and important research data in selected fields.

Conflicts over secrecy and openness in science are essentially conflicts over values. In order to explore the fundamental values which promote secrecy or openness in science, the American Association for the Advancement of Science has initiated a new project through the office of the AAAS Committee on Scientific Freedom and Responsibility. The project, titled "Secrecy and Openness in Scientific and Technical Communication" is supported by a grant from the Program on Ethics and Values in Science and Technology (EVIST) in the National Science Foundation, and the Humanities, Science and Technology Program in the National Endowment for the Humanities. Ms. Rosemary Chalk, Program Head for the AAAS Committee, is the project director.

The tradition of openness in research is the foundation for objectivity in science. It is through the free exchange of information and data that new ideas and experimental results are subjected to the rigorous test of peer review and verification. The origins of openness, however, have their roots in a period when science was essentially a private intellectual activity. Also, many scholars are not completely "open" in their exchange of data and information. Self-imposed restrictions on the release of new but unconfirmed theories or preliminary experimental data are quite common in traditional scientific work. These restrictions, which form part of the ethos of science,

are themselves limited by notions of fair play and equity, however, and are subject to abuse when stimulated by objectivies other than the protection of incomplete work.

In modern times, government, industrial and university groups have increasingly recognized the importance of applying scientific and technical resources to selected public and private objectives. Access to new information, including basic research, has emerged as a source of competitive advantage in the pursuit of various social, military, and economic goals. As a result, the concept of intellectual property has expanded in the post-World War II period to justify occasional controls on the disclosure of basic research findings supported by public or private funds.

For example, in a series of reports describing concerns about technology transfer leaking advanced U.S. technology to foreign adversaries, the Defense Department has questioned whether the openness associated with university research in areas of direct military application is detrimental to national security interests in a time of escalating East-West tensions.

In the commercial area, a number of firms are exploring arrangements whereby universities can develop research projects and academic programs suited to the needs of particular industries. Within such arrangements, one major source of concern and controversy is pre-publication review of, and patent protection for, new research data resulting from industry-sponsored work.

Secrecy also results from actions within the scientific community. As personal prestige, professional advancement and financial gains become more closely tied to publication, some individual scientists have indicated reluctance to exchange new research findings or materials with colleagues and students in the traditional manner.

These public and private pressures foster secrecy and science. Such restrictions on communication often serve legitimate and important social purposes. They may at times also result in arbitrary or abusive practices, or promote bias and the loss of objectivity in research.

Although there is reason to believe that secrecy is increasing in science, and that it may affect values other than openness, very little is known about the ways in which secrecy or openness influence the conduct of scientific research. It is for the purpose of encouraging attention to such relationships, and the values which affect professional behavior and education, that the AAAS Committee on Scientific Freedom and Responsibility has initiated the new project.

The AAAS project will consist of a series of background papers and regional seminars to be organized in 1984. Ten background papers will be commissioned through the project. Five projects seminars will be held in Boston, and one each will be held in Chicago, Nashville, San Diego and Washington, D.C. A project symposium will also be held as part of the 1984 AAAS Annual Meeting in New York.

Co-sponsoring institutions are:

American Association for the Advancement of Science, Committee on Scientific Freedom and Responsibility

Center for the Study of Ethics in the Professions, Illinois Institute of Technology (CSEP/IIT)

Management of Technology Programs, Vanderbilt University

Program in Science, Technology and Society, Massachusetts Institute of Technology (MIT)

Science, Technology and Public Affairs Program, University of California, San Diego (UCSD)

Science, Technology and Human Values

Regional hosts for the project are: Rosemary Chalk, AAAS project director, Washington, D.C.; Robert House, director, Management of Technology Program, Vanderbilt University; Marcel La Follette, editor, Science, Technology and Human Values, MIT; Sanford Lakoff, professor of political science, UCSD; and Vivien Weil, senior research associate, CSEP/IIT.

Advisory committee members guiding the development of the AAAS project are: Loren Graham, professor of the history of science, MIT; Harold P. Green, professor of law, George Washington University; Lee Grodzins, professor of physics, MIT; Louis Menand, senior lecturer in political science and special assistant to the provost, MIT; and Eugene Skolnikoff, director of MIT Center for International Studies.

Further information about the project can be obtained from Rosemary Chalk at American Association for the Advancement of Science, 1515 Massachusetts Ave. NW, Washington DC 20005 or call (202) 467-5238.

HAND-HELD CRYPTO DEVICE SEC-36

LOUIS KRUH

Tadiran Israel Electronics Industries Ltd., Tel Aviv, Israel, manufactures a line of communications security terminals and devices including the Digital Data Crypto Device SEC-15, which provides ciphering/deciphering of digital information for transmission over terminal-to-terminal radio or wire links; Secure Communication Terminal SEC-22, a state of the art digital narrow-band voice ciphering terminal; Secure Communication Terminal SEC-13, a digital ciphering/deciphering system designed to provide maximum security for tactical and data communication; and Hand-Held Crypto Device SEC-36, which this article will review.

Hand-Held Crypto Device SEC-36

DESCRIPTION

SEC-36 is a hand-held, battery powered military crypto device intended basically for tactical applications. The unit is operated off-line and provides encryption of messages for secure transmission over radio or wire communication channels.

The SEC-36 software incorporates an algorithm which encrypts the plaintext message in conjunction with a randomly generated key code supplied by the operator. The totally random code makes the ciphertext produced theoretically unbreakable, even when unauthorized persons gain knowledge of or access to the equipment. Decryption is possible only when the SEC-36 units at either end of the communication channel are set to the same algorithm and the same key code.

The SEC-36 features a five-character LED display and a push button keyboard with full alphanumerics and various function keys. The 3-level memory has a capacity of 90 five-character groups. Full editing facilities allow any part of the message to be recalled for review or amendment from any of the memory's levels (plaintext, code, and ciphertext).

OPERATION

Encryption

The encryption process involves all three levels of the memory:

PLAIN level
CODE level
CRYPTO level

Operate the PLAIN level key and type in the plaintext message in groups of five characters, using the alphanumeric keys. The LE-NO key selects alphabetic or numeric characters; the BLK key must be operated for each blank space in a group.

As the group is keyed in, it appears in the display. A dot above the first (left) character indicates that the memory's plain level is activated. A short beep is heard each time a five-character group is typed in and no more data can be entered. When a group is completed, operate the ENTR key to enter the group into the memory on plain level and clear the display for the next group. Repeat until end of message.

Operate the CODE level key and use the alphanumeric keys to type in the code sequence similarly in groups of five characters. (To ensure ciphering of the entire text, one code character should be entered for each plaintext character. The crypto level key is therefore inactivated until the correct number of code groups has been entered.) In the display, a dot above the middle character indicates that the code level in activated. After completing each group, press the ENTR key to enter the group into the memory on code level and clear the display.

Press the CRYPTO level key. The first group of ciphertext will appear in the display. A dot above the last (right) character indicates that the read-out comes from the memory's crypto level. Operate the ENTR key to display the following groups of ciphertext one by one. After noting down the ciphertext, the written enciphered message is ready for tape punching or transmission over the communication channel.

If the same message is to be transmitted to several recipients, security requires that for each recipient a different code sequence be used. The CLEAR CODE key -- pressed twice for protection -- will clear the memory's code level without affecting the basic message stored on plain level. This permits entering new code sequences and obtaining a different ciphertext of the message for each addressee.

Decryption

The procedure for decryption is the same as that for encryption:

Press the PLAIN level key and enter the received ciphertext by means of the alphanumeric keys. Operate the ENTR key after each group.

Press the CODE level key and enter the code sequence (obtained from an external source according to the identification group communicated by station of origin). Operate the ENTR key after each group.

Press the CRYPTO level key. The display will show the first group of decrypted plaintext. Operate the ENTR key to obtain the following groups one by one.

For editing or other purposes, the unit permits the user to recall to the display the plaintext, ciphertext or code sequence. Direct random access to any group stored in the memory is also available by selecting the appropriate level (level keys), operating the NUM key, and keying the desired group's address number. Corrections are made by using the ERS key which erases the contents of the display and permits a corrected group to be entered instead.

In response to specific questions, the company has said that the algorithm used in the SEC-36 is a Vigenere type, and the degree of security is dependent on the one-time key; the software which incorporates the algorithm is changeable; the unit is aimed at applications where the users need simple, unsophisticated means for multi-task encryption; and the price, for reasonable quantities, is about $5,000 per unit, with military specifications.

CIPHER EQUIPMENT

LOUIS KRUH

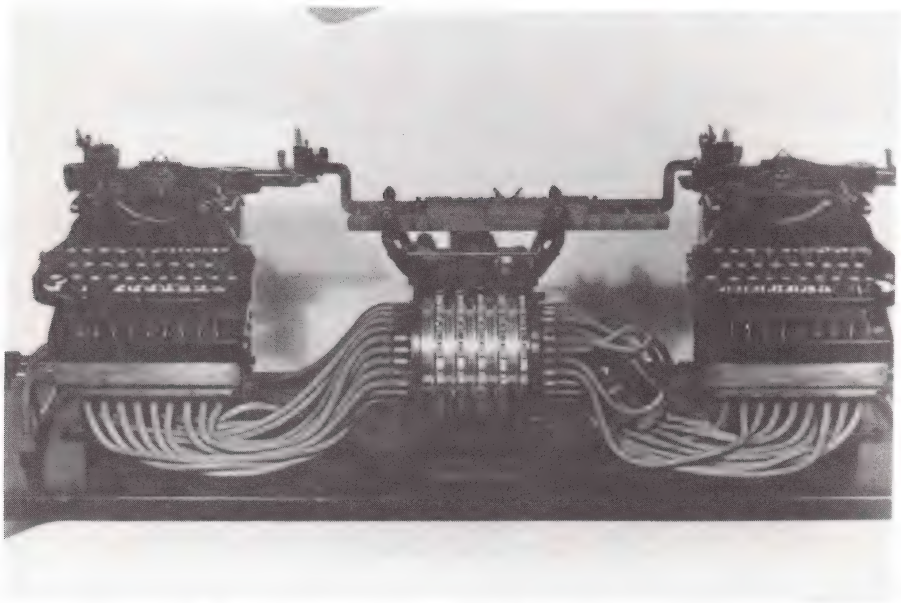
The Sidney Hole Cryptographic Machine, located in the Science Museum in London, is an impressive looking mechanism and unusual in that it uses compressed air in its operation.

Two of the machines wee manufactured in 1926 for the British War Office at a cost of 375 pounds each.

Our thanks to Donald W. Davies for a fine detailed description of a fascinating cryptographic device.

SIDNEY HOLE'S CRYPTOGRAPHIC MACHINE

DONALD W. DAVIES



Sidney Hole's Cryptographic Machine

This machine was patented in Britain in 1926 (Nos 207, 257, 239, 341 and 271, 955) and then in many other countries, but only one patent seems to have been obtained in USA (#1,684,028 in 1928) and Canada. The details here, and the drawings are from the Australian patent.

Two machines are held by the Science Museum and were inspected at their Hayes store. One seems to be an early development model and made rather crudely. The other is substantially the same as the drawings in the patent. In a parcel with the machine were many detailed drawings, many patent documents and a single page from the firm which made the model, which gives a little fragmentary history.

The better machine was examined. Almost none of its mechanism would move, due to corrosion though the parts looked undamaged. The rotors (see below) would rotate and some parts moved stiffly, otherwise nothing moved.

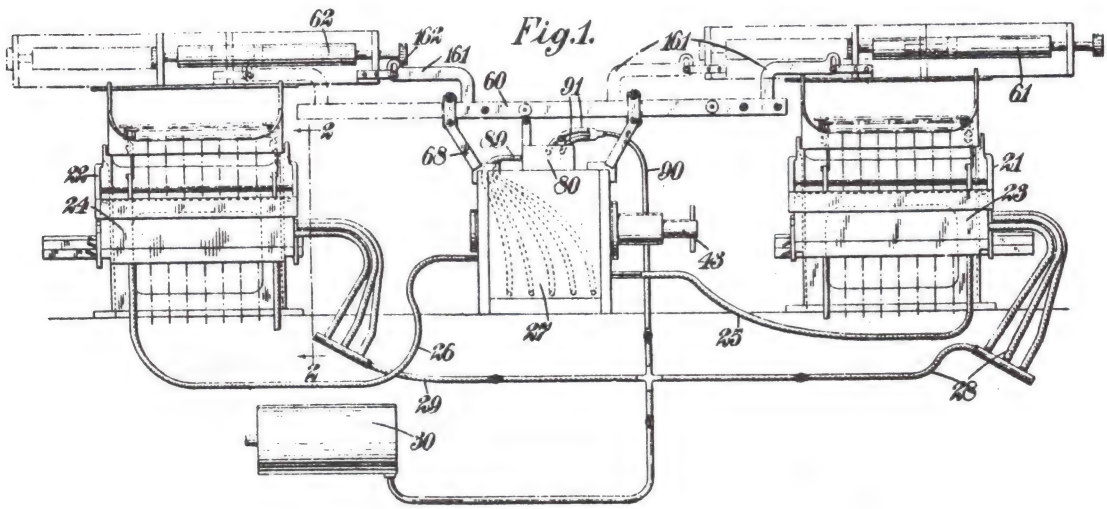


Figure 1.

The photograph and Figure 1 show the layout. Two modified typewriters are connected via a pneumatic rotor machine. The 28 keys of typewriters can either operate valves to apply suction to a pipe or receive suction from that pipe to actuate the key. Each of these functions is chosen by moving a slide composed of 3 bars passing through all the valve/piston assemblies, see Figure 2. By changing over the functions, encipherment or decipherment is performed, for example encipherment by pressing keys on the left typewriter to actuate

keys on the right via the rotor machine and decipherment by pressing keys on the right to actuate keys on the left. Both plain and cipher are printed in each case by the typewriter mechanisms.

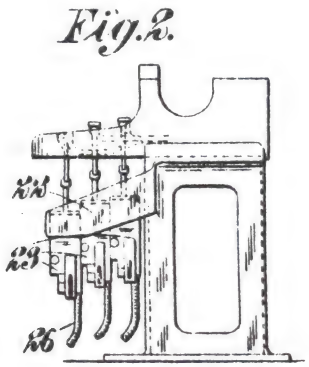


Figure 2.

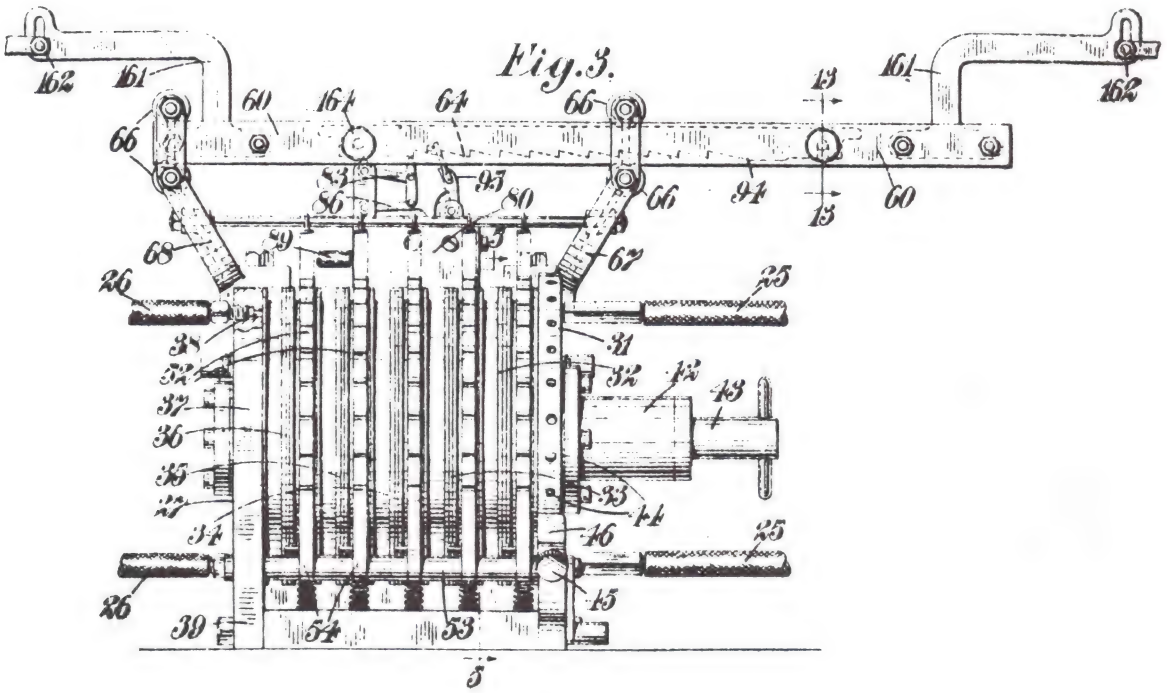


Figure 3.

The space bar has been removed. Spaces are shown by full stops in the text in the photograph. The shift mechanism is available but set manually on each machine separately. The carriages are linked mechanically to move in step. How the escapement of one is taken out of action to allow the other to drive was not clear.

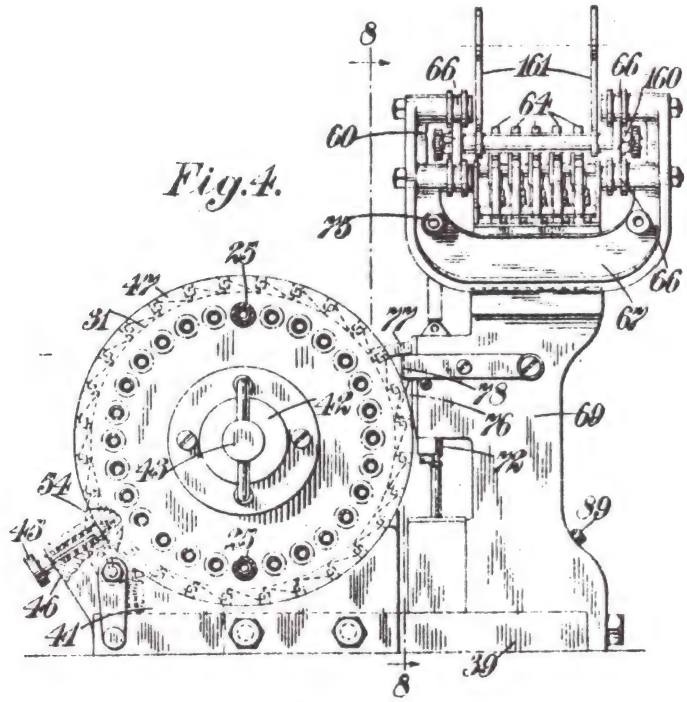


Figure 4.

Figure 3-7 show the rotor mechanism. It has 5 moving cylinders, each with 28 holes through it. Each one permutes the vacuum connection. 28 rubber pipes connect the ends to the typewriters.

A cylinder or rotor is made from 3 parts (see Figures 6 and 7). The outer parts carry grooves connected to the holes on the faces. The inner part is a thin plate with 28 holes to connect the grooves together. Like a 2 layer P. C. board this allows, in principle, any permutation of connections, but those actually used were rather limited (if less regular than the one shown by Figure 5). Like any rotor machine the permutation of letters is changed by rotating the cylinders, using the ratchet on one of their parts. The ratchet teeth are labelled A-Z. (number) where the number identifies the rotor type. By this means, initial settings can be listed.

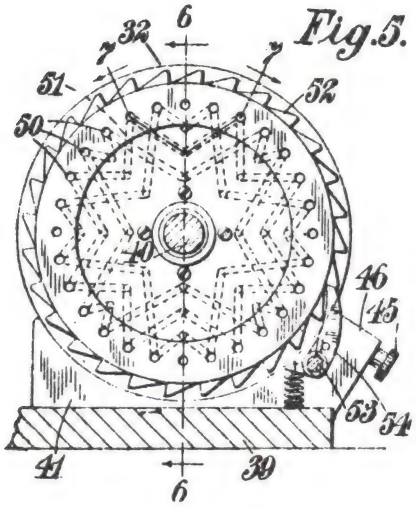


Figure 5.

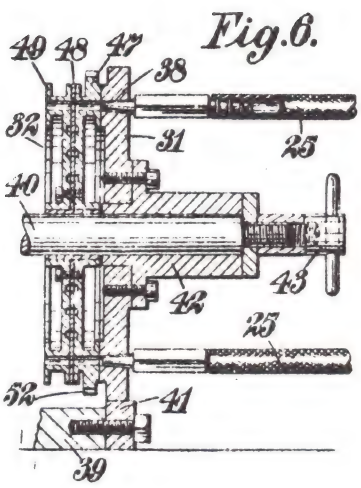


Figure 6.

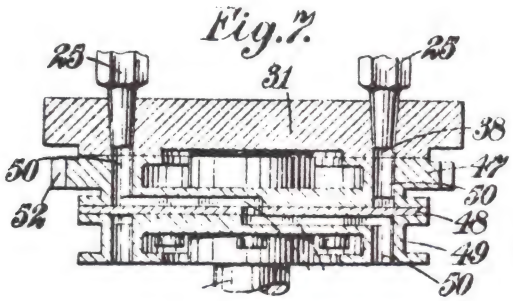


Figure 7.

In principle the pipes can be moved to new places. Their fit is a tapered plug, now corroded and fixed. Also the right hand end plate can be rotated if pin 45 is pulled but that is also stuck. By removing nut 43 (stuck) the rotor mechanism could be disassembled and the rotor changed. No spare rotor cylinders were provided.

The movement of the rotors is determined by carriage movement. Hence, a particular rotor steps on at a particular column on the printed messages. It will be seen that complete carriage movements are enforced, and the fixed sequence of rotor steps is the same on each line of message.

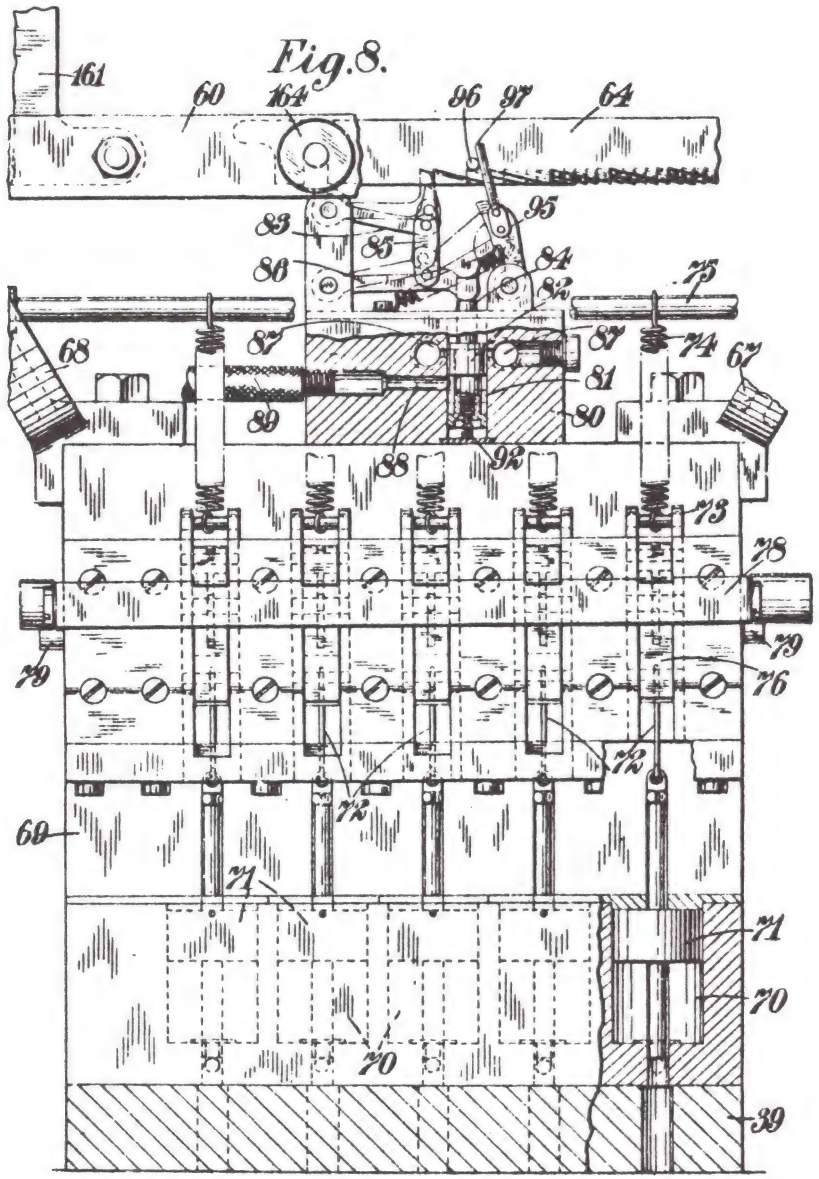


Figure 8.

The frame linking the carriages holds 5 toothed racks shown as 64 in Figures 3, 4, 8, 9, and 11. Each rack actuates a follower 83 (Figures 8, 11) which moves a valve 82 (Figures 8, 12) which applies suction to a cylinder 70

(Figures 8, 9) with a piston 71 to move a pawl 76 (Figure 9) to step a rotor cylinder. There are five of thee followers, valves, pistons and pawls. The valves are connected to the cylinders by five rubber pipes 89. The vacuum supply is 91 (two pipes). The return of each piston is by a spring, sucking air back, though where from is not clear.

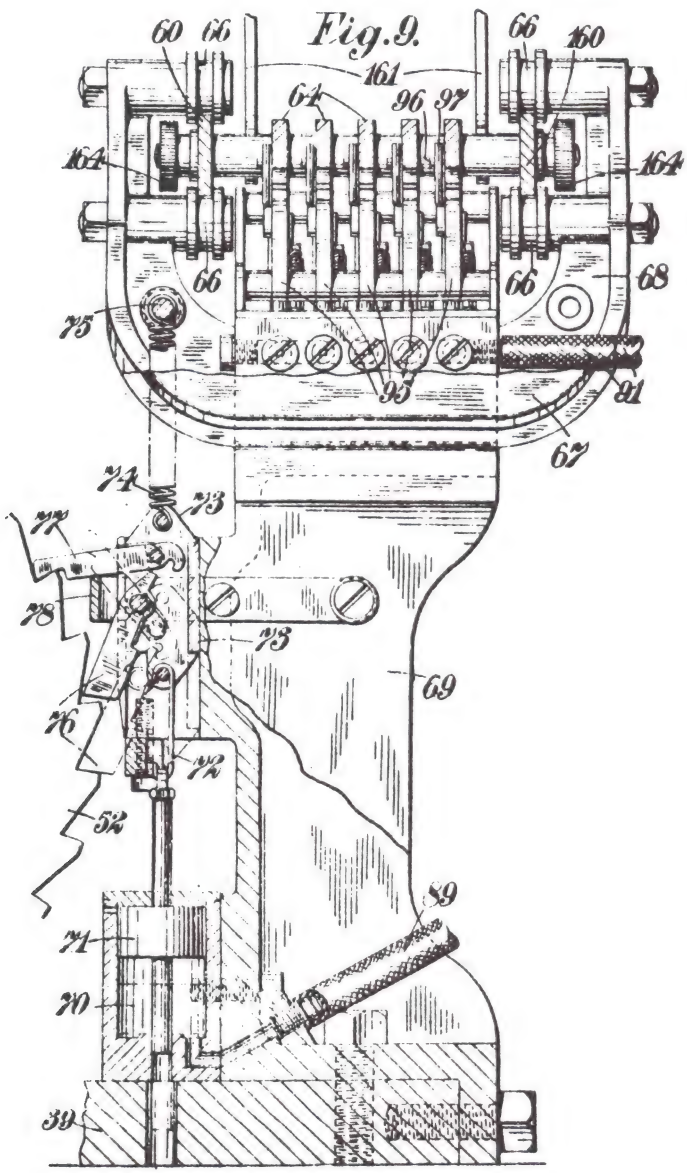


Figure 9.

The racks have ratchet type teeth. It is the upward movement of the follower over the edge which triggers the stepping pawl. The return of the pawl is gradual, due to the slopes of the rack teeth, but this is immaterial since the rotor never moves more often than once in 5 characters.

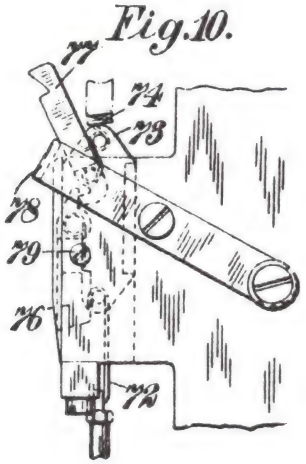


Figure 10.

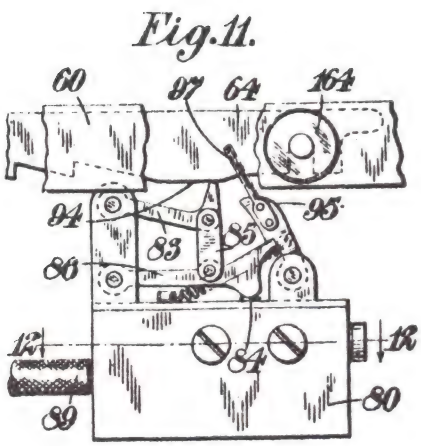


Figure 11.

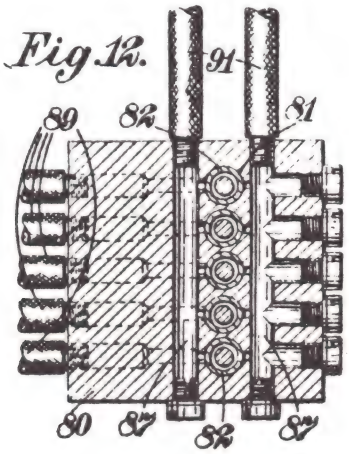


Figure 12.

Pawls 54 (Figure 5) prevent reverse motion of the rotors but can be withdrawn by a lever to allow initial setting. Pawls 77 (Figure 9 and 10) probably prevent forward motion when the piston is returned. They withdraw as the pistons operate, see broken lines in Figure 9. An arm 78 withdraws these pawls for initial setting.

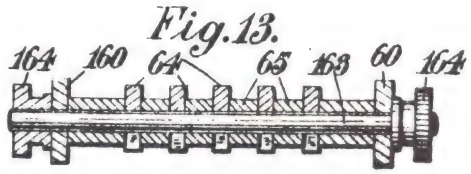


Figure 13.

The teeth of the racks would catch on the followers to prevent the carriages returning, and this does prevent premature carriage return. At the end of travel, cams 94 (Figures 3 and 11) press the followers down beyond their normal travel and allow detents 95 (Figure 11) to catch them. They remain in this position until the carriage is completely returned when pins 96 on each rack (Figure 8) knock off the detents and let the followers fall back into action. This mechanism ensures that the cycle of rotor stepping is the same on each carriage movement and that each carriage movement is complete. In the photograph it seems that the lines of message have been completed to the end by full stops ... etc.

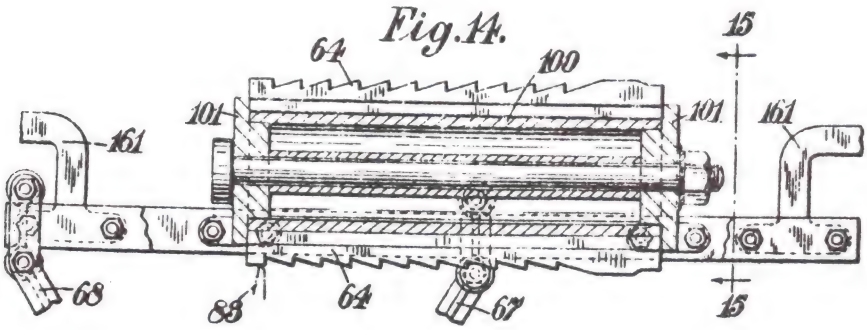


Figure 14.

The five racks have a pattern of teeth which is part of the key to the cipher. They were removable be slackening the knurled nuts 164 (Figure 8, 11). Each rack on the model inspected was double-sided. The inventor suggests that a set of racks be used as a physical key to transport to the authorized cipher

machine. He also suggests different methods of constructing this control element - a 'barrel of racks' in Figures 14, 15, 16 or a rotary controller in Figures 17, 18 or a pneumatic rotary controller in Figures 19, 20.

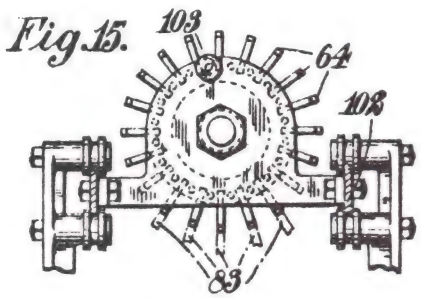


Figure 15.

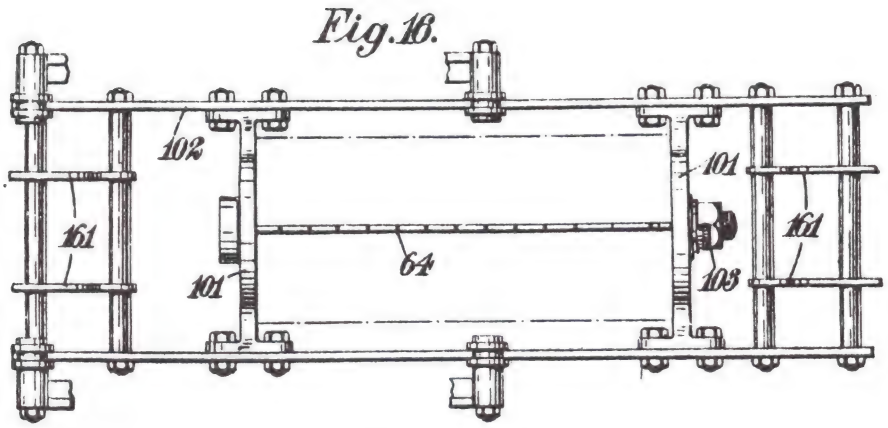


Figure 16.

Given the five rack system examined, arbitrary use of the 10 side is not suitable if we require that at least one rotor should move at each carriage movement. The actual racks show great regularity. Mostly, the five rotors move in sequence, except that one rack in use had some steps missing and another had missing steps on its upper, unused side. The patterns were not recorded. In particular, the number of rotor steps per carriage movement on each rack should be prime to 28. The only evidence we have is in Figures 3 and 14 which show 11 steps in each case. The possibility of stepping more than one rotor at a time does not seem to have been exploited. After 28 lines of type have been printed, all rotors will have returned to their initial positions - a considerable weakness.

In its present state there is no way to decide how reliable the machine would be. The carriage movement probably took place as the typewriter key was released. This movement triggered the rotor movements. Was the carriage movement sufficiently precise? Did the rotors move fast, so that another key could be depressed soon? It is probable that the strokes must be made deliberately, as in a teleprinter.

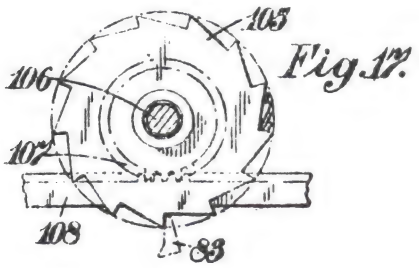


Figure 17.

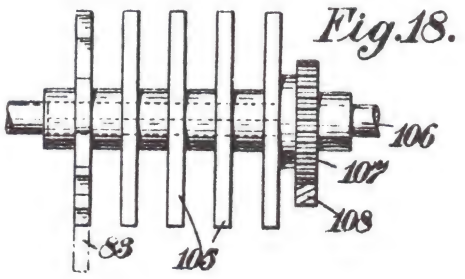


Figure 18.

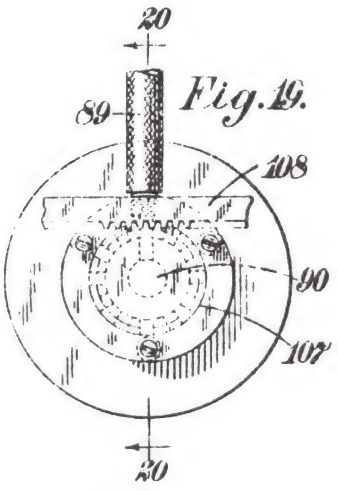


Figure 19.

It is possible that the machine could be restored to operation, by careful work. The detailed design and construction was by the Accounting and Tabulating Corporation of Great Britain Limited, later absorbed, as Power-Samas, into ICT and subsequently ICL. Its construction seems to be solid and precise. The bulk and weight and the need of a vacuum supply were the main drawbacks.

LITERATURE REVIEWS

LOUIS KRUH

KRUH ON KAHN ON CODES

Kahn, David. Kahn on Codes: Secrets of the New Cryptography. Macmillan Pub. Co., 866 Third Ave., New York NY 10022. 1983. 343 pp. $19.95.

As an author, David Kahn is usually associated with his first book, that massive tome which has become a classic, The Codebreakers.

His many articles, whether they appear in scholarly publications such as Aerospace Historian, The Historical Journal or Isis; technical journals like Cryptologia or Computers and Security; society journals like The Cryptogram; intellectual outlets such as The New York Times Book Review or The New Republic; establishment journals like Foreign Affairs; or initially delivered as a keynote address at an international historical conference; or as a statement to a United States Congressional Committee; individually appear insignificant compared to that weighty, encyclopedic work.

But, when 28 of his best articles are assembled in one volume, it not only casts its own shadow, it also disproves the axiom which claims, "the whole is only equal to the sum of its parts." There are at least three reasons for this. The reprinted stories as a group are better than their original versions because they are updated or corrected as later events have dictated; many of the articles which appeared in publications whose format did not allow for footnotes, or were delivered as talks, have been carefully documented with extensive references; and, of course, there is the convenience of having them all in one volume.

This excellent collection is arranged in seven sections: "Uncovering Cryptology's Past," "Overviews," "Historical and Technical Studies," "The Politics of Cryptology," "Book Reviews," "Codes in Context," and "The Future."

The articles include stories of Kahn's meeting with famous European cryptologists, an analysis of codebreaking in World War I and World War II, a searching examination of NSA, a study of military intelligence in action, opportunities for further historical research in cryptology, and much more.

One article was written expressly for this book, "The Spy Who Most Affected World War II." For that distinction, Kahn designates Hans-Thilo Schmidt, an obscure Nazi party member, who was a civilian clerk in the German Signal Corps. Schmidt was the spy who delivered documents that Polish cryptanalysts used to solve the German Enigma cipher machines. This article reveals for the first time the background and details of this virtually unknown man and his unparalleled betrayal which had such an enormous impact on the outcome of World War II.

Besides being a comprehensive and exciting account of notable cryptologic events, the book is a genuine bargain with its price only a fifth of what it would cost to buy copies of the publications in which the original articles appeared.

"INSIDE" NSA?

Bamford, J. The Puzzle Palace: A Report on America's Most Secret Agency, Houghton Mifflin Co., 2 Park St., Boston MA 02108, 1982, 465 pp., $16.95

This is the most comprehensive book ever written about the National Security Agency and it contains an amazing amount of detail starting from its inception as MI-8 in WW I to the present day. Its origin is traced through H. O. Yardley, W. F. Friedman, Pearl Harbor, WW II, and the various studies/committee investigations on unification of cryptologic activities, which ultimately led to President Truman's still secret 1952 memorandum establishing NSA. Bamford describes NSA's Fort Meade headquarters -- he refers to it as SIGINT City — the physical layout and organization, how it operates, its worldwide influence, and many of its senior officials. President John F. Kennedy once told the intelligence community, "Your successes are unheralded; your failures are trumpeted." As if to underscore the truth of that remark, Bamford is only able to relate few of NSA's triumphs but, almost with excessive zeal, reveals all of its warts, and virtually all are twice-told tales. Where the author has found new information, particularly dealing with personalities, as in most of the chapter on cooperation between the British GCHQ and NSA, it makes for interesting reading. On the other hand, the section on NSA's complex network of listening posts with details on antennas, circuits, microwave signals and locations of secret sites, which is the book's largest chapter and contains new data, will undoubtedly be dull to many readers except for those inimical to NSA's mission.

A great deal of information was derived from an assiduous study of NSA's almost 30-year old, unclassified 20-page monthly newsletter, which the author wrangled from the Agency, from extensive research among the Friedman Papers at

the George C. Marshall Research Library, and many interviews with former NSA officials. The overall result is a fascinating glimpse at previously unpublished items about Friedman, Callimahos, a host of other lesser known key officials, and the intriguing life and times in SIGINT City.

INTELLIGENCE BIBLIOGRAPHY

Constantinides, G.C. Intelligence and Espionage: An Analytical Bibliography. Westview Press, 5500 Central Ave., Boulder, CO 80301, 1983, 559 pp., $60.

The author, who has spent almost 25 years in U.S. government intelligence and national security work, has justifiably described his book as "the most comprehensive and thorough bibliography of English-language nonfiction books on intelligence and espionage to date." It is an enormous work with knowledgeable comments, most of them a page or more, on close to 500 books. In a special category index the author has divided them into 54 categories. The bibliography itself is arranged by author. One of the categories is Communications Intelligence, Cryptology, and Signals Intelligence which contains 40 books. Constantinides demonstrates a familiarity and expertise in the subject matter with incisive comments and cross references in many of his annotations. In his remarks on Yardley's American Black Chamber, he provides views from five other authors and suggests areas in Yardley's career which still need to be explained. With Lewin's Ultra Goes To War, he refers to reviewers of the book as well as other authors to point out inaccuracies and to remind us that because much Ultra material is still secret, the full story has not yet been told. In his overall excellent appraisal of The Codebreakers, he expresses possibly an insider's view that Kahn's assessment of Friedman as being responsible for the U.S.' cryptologic superiority is questionable. Other worthwhile comments abound in this outstanding reference work which will be consulted frequently by persons seeking a guide to intelligence literature.

YARDLEY'S CHINESE BLACK CHAMBER

Yardley, H.O. The Chinese Black Chamber: An Adventure in Espionage. Houghton Mifflin, 52 Vanderbilt Ave., New York, NY 10017, 1983, 225 pp., $13.95

In 1938, Chiang Kai-shek, head of the Nationalist Chinese government which was fighting a desperate losing battle against the Japanese, engaged Yardley to come to the war torn capital of Chungking to set up a Chinese version of the American Black Chamber Yardley had organized and directed in New York. This manuscript, hidden for over 40 years, is the story of his adventures and

intelligence exploits in China from 1938-1940. Most of the account is a fascinating glimpse of life in a strange society of Chinese characters, European traders, politicians, generals, spies, traitors, mistresses and other colorful personalities. Few of Yardley's cryptanalytical episodes are included but he does describe, step-by-step, how he solved a cipher which used a public Chinese code book superenciphered by a book cipher. The book has an introduction by James Bamford, author of The Puzzle Palace, with additional details of Yardley's experiences in China. It concludes with "Memories of the American Black Chamber", a brief memoir by the author's wife, Edna Yardley, who is its last surviving original member.

CODES IN THE ETHER

Monitoring Times, 140 Dog Branch Road, Brasstown NC 28902. Issued monthly, 32 pp., $10.50 for one year, $20.00 for two years.

This 32 page tabloid newspaper is written for shortwave listeners and scanner buffs. It usually has a feature on clandestine stations including spy number broadcasts, i.e., stations transmitting messages in numerical code. A recent issue contained articles on Basic Codebreaking and Japanese messages sent before the Pearl Harbor attack. It covers other offbeat listening areas such as satellite reception, monitoring the AWACS Net, nuclear shipments, etc. and provides the frequency lists. Other features review equipment, books, provide advice on getting started and improving your operation. Free sample copy available on request.

BIOGRAPHY OF ALAN TURING

Hodges, A. Alan Turing: The Enigma. Simon and Schuster, 1230 Ave. of the Americas, New York NY 10020. 1983. 587 pp. $22.50.

Alan M. Turing was an English mathematical genius whose name is perpetuated in the annals of computer history for the Turing machine, a theory and, eventually, a device he invented. His work, starting in the mid-1930s, was the theoretical foundation for the modern digital computer.

Lesser known is his leading role at Bletchley Park, where Government Code and Cypher School cryptanalysts were confronted with improved Enigma ciphers when the Germans upgraded their communications security. Instead of using six or seven plugboard connections, the Enigma operators started to connect ten pairs of letters; and the number of available rotors was increased from three to five. There are 150,738,274,937,250 ways to connect ten pairs of letters, and with just three rotors there are six ways to arrange them -- but with five to

choose from, the number of possibilities jumps to sixty! Polish cryptanalysts, who first broke the Enigma cipher and developed the Bombe, an electromechanical device to run through all possible rotor positions, simply didn't have the technical resources to continue, and their information was given to the British and the French.

Turing, with the assistance of Gordon Welchman, developed a new Bombe embodying a new concept and a new design which helped to decrypt the new Enigma system in speedy fashion.

Later, Turing was assigned to work on the German Naval Enigma which at that time still used three rotors but they were chosen from a group of eight which produced 336 possibilities. Turing's analysis demonstrated that decryptment would be impossible until additional enciphering information could be captured. In the interim, he developed the mathematical theory that was required to exploit the information when it became available.

When teleprinter-enciphering machine traffic known as "fish" was being analyzed, one of the most important and general methods was invented by Turing and became known as Turingismus.

This detailed biography covers Turing's life from birth to his untimely death in an intelligent, exhaustive and very readable style. Because the author is also a mathematician, he is able to explain and make comprehensible the technical passages dealing with Turing's cryptanalytical and mathematical achievements. It is the kind of biography you wish would be done for William F. Friedman.

ON KULLBACK'S χ-TESTS FOR MATCHING AND NON-MATCHING MULTINOMIAL DISTRIBUTIONS

BORGE TILT

ABSTRACT: This paper concerns the χ-tests (cross-product sum tests) introduced by Solomon Kullback in 1938 ("Statistical Methods in Cryptanalysis") to test for matching or non-matching of parameters of two multinomial distributions with known and identical sets of parameters, namely the probabilities. Our main result is the demonstration of an error in the given expression for the variance of χ in the case of non-matching distributions.

KEYWORDS: multinomial distribution, cross-product sum tests

Among the tools of the cryptanalyst is the cross-product sum test, or χ-test, devised by Solomon Kullback in 1935. Later, in 1938, it was described in his "Statistical Methods in Cryptanalysis," which, however, was not declassified until quite recently [1].

The probability distribution of the test variable χ will often be well approximated by a normal distribution, thus requiring knowledge of mean and variance only. Kullback gives formulas for mean and variance of χ in two cases: (1) matching distributions, and (2) non-matching distributions. Our main objective is to point out an error in the expression for the variance of χ in the case of non-matching distributions, and to derive the correct formula.

We give mean and variance of χ for both matching and non-matching distributions. We also derive the mean of χ in an intermediate case not discussed by Kullback, partially matched distributions.

THE CROSS-PRODUCT SUM

The distributions in question are two multinomial distributions with paramters $(N_1;\ p_1,p_2,\ldots,p_n)$ and $(N_2;\ \pi_1,\pi_2,\ldots,\pi_n)$, respectively, where (p_i) and (π_i) are permutations of the n probabilities in a set $\{h_k\}$ with $0 \leq h_k \leq 1$ for $k = 1,2,\ldots,n$ and $\Sigma_{k=1}^{n} h_k = 1$. The pairs (p_1,π_1), (p_2,π_2), ..., (p_n,π_n) are formed

according to a specification depending on the particular case, but the order of the pairs does not matter.

Suppose a single observation is made of each multinomial variable, and let $(a_1,a_2,\ldots,a_n)$ and $(b_1,b_2,\ldots,b_n)$ denote the results. Thus, for instance, a_i $(i = 1,2,\ldots,n)$ is the number of times outcome number i has occurred in N_1 independent trials, where outcome number i $(i = 1,2,\ldots,n)$ occurs with probability p_i. The cross-product sum is

$$\chi = \sum_{i=1}^{n} a_i b_i. \tag{1}$$

Instead of χ we will sometimes consider the variable

$$\chi/N_1N_2 = \sum_{i=1}^{n} (a_i/N_1)(b_i/N_2), \tag{2}$$

which is a cross-product sum of the relative frequencies.

We shall discuss the probability distribution of $\chi(\chi/N_1N_2)$ in three cases distinguished by the method of pairing p's and π's. Generally, some pairs of identical p and π, possibly none at all, are matched in advance, and the remaining p's and π's are paired at random.

MATCHING DISTRIBUTIONS

This is the case where all p's and π's are matched. Thus $p_i = \pi_i$ for all i. The ordering of p's (π's) is immaterial as far as the probability distribution of χ is concerned. In [1, p.52] it has been shown that mean and variance of χ are, respectively,

$$E[\chi] = N_1N_2s_2 , \tag{3}$$

$$V[\chi] = N_1N_2[(N_1+N_2)(s_3-s_2^2) + (s_2+s_2^2-2s_3)], \tag{4}$$

where $s_2 = \Sigma_{k=1}^{n} h_k^2$ and $s_3 = \Sigma_{k=1}^{n} h_k^3$.

Always, $s_3 - s_2^2 \geq 0$, and $s_2 + s_2^2 - 2s_3 \geq 0$.

By (3) and (4),

$$E[\chi/N_1N_2] = s_2, \tag{5}$$

$$V[\chi/N_1N_2] = (1/N_1 + 1/N_2)(s_3-s_2^2) + (s_2+s_2^2-2s_3)/N_1N_2. \tag{6}$$

Thus, whereas $E[\chi/N_1N_2]$ is a constant, independent of N_1 and N_2, $V[\chi/N_1N_2]$ is a decreasing function of N_1 and N_2. In fact,

$$V[\chi/N_1N_2] \longrightarrow 0 \text{ as } N_1, N_2 \longrightarrow \infty. \tag{7}$$

NON-MATCHING DISTRIBUTIONS

Here, none of the p's and π's are matched, so all p's and π's are paired at random. We shall prove

$$E[\chi] = N_1N_2 \cdot \frac{1}{n}, \tag{8}$$

$$V[\chi] = N_1N_2 \frac{[N_1(s_2 - \frac{1}{n}) + (1-s_2)]\cdot[N_2(s_2 - \frac{1}{n}) + (1-s_2)]}{n-1}, \tag{9}$$

where $s_2 - 1/n$ and $1-s_2 \geq 0$. Equivalently,

$$E[\chi/N_1N_2] = \frac{1}{n}, \tag{10}$$

$$V[\chi/N_1N_2] = \frac{[(s_2-\frac{1}{n}) + (1-s_2)/N_1]\cdot[(s_2 - \frac{1}{n}) + (1-s_2)/N_2]}{n-1}. \tag{11}$$

Again, $E[\chi/N_1N_2]$ is a constant, independent of N_1 and N_2, and $V[\chi/N_1N_2]$ is decreasing in N_1 and N_2. However, $V[\chi/N_1N_2]$ does not, in general, converge to 0, since

$$V[\chi/N_1N_2] \longrightarrow \frac{(s_2 - \frac{1}{n})^2}{n-1} \quad \text{as } N_1, N_2 \longrightarrow \infty \tag{12}$$

It is worth noting that in the special case of a flat distribution, that is $h_k = 1/n$ for all k, it makes no difference whether p's and π's are matched or not. In this very special case $s_2 = 1/n$ and $s_3 = 1/n^2$. Substitution of these

values into both Eq.(3) and Eq.(8) gives $E[\chi] = N_1N_2 \cdot \frac{1}{n}$, and substitution into both Eq.(4) and Eq.(9) gives $V[\chi] = N_1N_2 \cdot \frac{1}{n}(1 - \frac{1}{n})$.

Kullback's formula for $V[\chi]$ in the case of non-matching distributions is not in agreement with our Eq. (9). According to Kullback [1, p.53],

$$V[\chi] = N_1N_2[(N_1+N_2)(\frac{s_2}{n} - \frac{1}{n^2}) + \frac{1}{n} + \frac{1}{2} - \frac{2s_2}{n_n}], \qquad \text{(Kullback, Eq.(21.7))}$$

by which $V[\chi/N_1N_2] \to 0$ as $N_1, N_2 \to \infty$. Only if $N_1 = 1$ or $N_2 = 1$, or if $h_k = 1/n$ for all k, does the above formula give the same result as our Eq.(9). In any other case it will underestimate the true value of the variance.

Method of Proof

For the purpose of proving Eqs.(8) and (9), we accomplish the required random pairing of p's and π's by letting $(p_1,p_2,\ldots,p_n)$ and $(\pi_1,\pi_2,\ldots,\pi_n)$ be two independent, random permutations of the probabilities in the set $\{h_k\}$.

We shall derive $E[\chi]$ and $E[\chi^2]$ which give us both $E[\chi]$ and $V[\chi] = E[\chi^2] - (E[\chi])^2$. The line of proof for $E[\chi]$ as well as $E[\chi^2]$ is: first find the conditional mean given fixed permutations $(p_1,p_2,\ldots,p_n)$ and $(\pi_1,\pi_2,\ldots,\pi_n)$, and next derive the unconditional mean by letting $(p_1,p_2,\ldots,p_n)$ and $(\pi_1,\pi_2,\ldots,\pi_n)$ be random permutations.

Derivation of $E[\chi]$

The conditional mean, given permutations (p_i) and (π_i), is

$$\begin{aligned} E[\chi \mid (p_i),(\pi_i)] &= E[\Sigma_{i=1}^n a_i b_i] \\ &= \Sigma_{i=1}^n E[a_i]E[b_i] \\ &= \Sigma_{i=1}^n (N_1p_i)(N_2\pi_i). \end{aligned}$$

Thus,

$$E[\chi \mid (p_i),(\pi_i)] = N_1N_2\Sigma_{i=1}^n p_i\pi_i. \qquad (13)$$

Unconditioning, we let (p_i) and (π_i) be independent, random permutations. Thus p_i and π_i are independent, with $E[p_i] = E[\pi_i] = \Sigma_{k=1}^n h_k \frac{1}{n} = \frac{1}{n}$, for all i. Hence,

$$\begin{aligned} E[\chi] &= E[E[\chi \mid (p_i),(\pi_i)]] \\ &= E[N_1N_2\Sigma_{i=1}^n p_i\pi_i] \\ &= N_1N_2\Sigma_{i=1}^n E[p_i]E[\pi_i]. \end{aligned}$$

Thus,

$$E[\chi] = N_1 N_2 \frac{1}{n}. \tag{8}$$

We remark that Eq.(8) follows easily from the observation that χ is the total number of identical outcomes among the N_1N_2 pairs of trials, and that, given random pairing of p's and π's, the probability of identical outcomes is 1/n.

Derivation of $E[\chi^2]$

Again, we begin by deriving the conditional mean, given permutations (p_i) and (π_i).

$$\begin{aligned} E[\chi^2 \mid (p_i),(\pi_i)] &= E[\Sigma_{i=1}^{n} a_i b_i)^2] \\ &= E[\Sigma_{i=1}^{n} a_i^2 b_i^2 + 2\Sigma_{i=1}^{n-1}\Sigma_{j=i+1}^{n} a_i a_j b_i b_j] \\ &= \Sigma_{i=1}^{n} E[a_i^2]E[b_i^2] + 2\Sigma_{i=1}^{n-1}\Sigma_{j=i+1}^{n} E[a_i a_j]E[b_i b_j]. \end{aligned}$$

Now, a_i is a binomial variable with mean N_1p_i and variance $N_1p_i(1-p_i)$, so that $E[a_i^2] = N_1p_i(1-p_i) + (N_1p_i)^2$. Furthermore, $E[a_ia_j] = N_1(N_1-1)p_ip_j$. The last formula is derived as follows: $E[a_ia_j|a_i] = a_iE[a_j|a_i] = a_i(N_1-a_i)p_j/(1-p_i)$, $E[a_ia_j] = E[E[a_ia_j|a_i]] = (N_1E[a_i] - E[a_i^2])p_j/(1-p_i) = N_1(N_1-1)p_ip_j$, using the expressions for $E[a_i]$ and $E[a_i^2]$. Similarly, $E[b_i^2] = N_2\pi_i(1-\pi_i) + (N_2\pi_i)^2$ and $E[b_ib_j] = N_2(N_2-1)\pi_i\pi_j$. By insertion of these expressions we obtain

$$\begin{aligned} E[\chi^2 \mid (p_i),(\pi_i)] = {} & \Sigma_{i=1}^{n}(N_1(N_1-1)p_i^2 + N_1p_i)(N_2(N_2-1)\pi_i^2 + N_2\pi_i) \\ & + 2\cdot\Sigma_{i=1}^{n-1}\Sigma_{j=i+1}^{n}(N_1(N_1-1)p_ip_j)(N_2(N_2-1)\pi_i\pi_j). \end{aligned}$$

A rearrangement gives

$$\begin{aligned} E[\chi^2 \mid (p_i),(\pi_i)] = {} & N_1N_2(N_1-1)(N_2-1)\Sigma_{i=1}^{n}p_i^2\pi_i^2 + N_1N_2\Sigma_{i=1}^{n}p_i\pi_i \\ & + N_1N_2(N_1-1)\Sigma_{i=1}^{n}p_i^2\pi_i + N_1N_2(N_2-1)\Sigma_{i=1}^{n}p_i\pi_i^2 \\ & + N_1N_2(N_1-1)(N_2-1)\cdot 2\cdot\Sigma_{i=1}^{n-1}\Sigma_{j=i+1}^{n}p_ip_j\pi_i\pi_j. \end{aligned} \tag{14}$$

Unconditioning, but utilizing only the independence between (p_i) and (π_i),

$$E[\chi^2] = E[E[\chi^2 \mid (p_i),(\pi_i)]]$$

$$= N_1N_2(N_1-1)(N_2-1)\Sigma_{i=1}^{n}E[p_i^2]E[\pi_i^2] + N_1N_2\Sigma_{i=1}^{n}E[p_i]E[\pi_i]$$

$$+ N_1N_2(N_1-1)\Sigma_{i=1}^{n}E[p_i^2]E[\pi_i] + N_1N_2(N_2-1)\Sigma_{i=1}^{n}E[p_i]E[\pi_i^2]$$

$$+ N_1N_2(N_1-1)N_2-1)\cdot 2\cdot\Sigma_{i=1}^{n-1}\Sigma_{j=i+1}^{n}E[p_ip_j]E[\pi_i\pi_j].$$

Since (p_i) and (π_i) are random permutations, we have

$$E[p_i] = E[\pi_i] = \Sigma_{k=1}^{n}h_k\,\frac{1}{n} = \frac{1}{n} \qquad (i=1,2,\ldots,n),$$

$$E[p_i^2] = E[\pi_i^2] = \Sigma_{k=1}^{n}h_k^2\,\frac{1}{n} = \frac{s_2}{n} \qquad (i=1,2,\ldots,n),$$

$$E[p_ip_j] = E[\pi_i\pi_j] = \Sigma_{k=1}^{n}\Sigma_{\substack{r=1\\ r\neq k}}^{n}h_kh_r/n(n-1)$$

$$= (\Sigma_{k=1}^{n}\Sigma_{r=1}^{n}h_kh_r - \Sigma_{k=1}^{n}h_k^2)/n(n-1)$$

$$= \frac{1-s_2}{n(n-1)} \qquad (i\neq j).$$

Substitution of these expressions for the means results in

$$E[\chi^2] = N_1N_2(N_1-1)(N_2-1)\,\frac{s_2^2}{n} + N_1N_2\,\frac{1}{n}$$

$$+ N_1N_2(N_1+N_2-2)\,\frac{s_2}{n}$$

$$+ N_1N_2(N_1-1)(N_2-1)\cdot 2\cdot\frac{n(n-1)}{2}\,\frac{(1-s_2)^2}{(n(n-1))^2},$$

where $n(n-1)/2$ is the number of terms in the double sum.

Thus

$$E[\chi^2] = N_1N_2\ [(N_1-1)(N_2-1)\ \frac{s_2^2}{n} + \frac{1}{n} + (N_1+N_2-2)\ \frac{s_2}{n} + (N_1-1)(N_2-1)\ \frac{(1-s_2)^2}{n(n-1)}]. \quad (15)$$

Derivation of $V[\chi]$

Insertion of the above expressions for $E[\chi]$ and $E[\chi^2]$ into the formula $V[\chi] = E[\chi^2] - (E[\chi])^2$, and collection of terms, give us

$$V[\chi] = N_1N_2\ [N_1n_2(\frac{s_2^2}{N} + \frac{(1-s_2)^2}{n(n-1)} - \frac{1}{n^2})$$

$$+ (N_1+N_2)(-\frac{s_2^2}{n} + \frac{s_2}{n} - \frac{(1-s_2)^2}{n(n-1)})$$

$$+ (\frac{s_2^2}{n} + \frac{1}{n} - \frac{2s_2}{n} + \frac{(1-s_2)^2}{n(n-1)})].$$

This reduces to

$$V[\chi] = N_1N_2\ [N_1N_2\ \frac{(s_2 - \frac{1}{n})^2}{n-1} + (N_1+N_2)\frac{(s_2 - \frac{1}{n})(1-s_2)}{n-1} + \frac{(1-s_2)^2}{n-1}]. \quad (16)$$

Finally, by factorization of Eq. (16) we obtain Eq. (9) above.

Variance decomposition

Observe that, by Eq. (16),

$$V[\chi/N_1N_2] = \frac{(s_2 - \frac{1}{n})^2}{n-1} + (\frac{1}{N_1} + \frac{1}{N_2})\frac{(s_2 - \frac{1}{n})(1-s_2)}{n-1} + \frac{1}{N_1N_2}\frac{(1-s_2)^2}{n-1}. \quad (17)$$

The variance $V[\chi/N_1N_2]$ may be split into two components due to random pairing and multinomial sampling variation, respectively. By application of the well-known general formula $V[X] = V[E[X|Y]] + E[V[X|Y]]$, see for instance [2,p.97], we conclude that

$$V[\chi/N_1N_2] = V[E[\chi/N_1N_2|(p_i),(\pi_i)]] + E[V[\chi/N_1N_2|(p_i),(\pi_i)]]. \quad (18)$$

The right-hand side of Eq. (18) is the sum of the variance of the conditional mean, given (p_i) and (π_i), and the mean of the conditional variance, given (p_i) and (π_i).

We shall show that

$$V[E[\chi/N_1N_2\,|\,(p_i),(\pi_i)]] = \frac{(s_2 - \frac{1}{n})^2}{n-1}\ , \tag{19}$$

$$E[V[\chi/N_1N_2\,|\,(p_i),(\pi_i)]] = (\frac{1}{N_1} + \frac{1}{N_2})\frac{(s_2 - \frac{1}{n})(1-s_2)}{n-1} + \frac{1}{N_1N_2}\,\frac{(1-s_2)^2}{n-1}\ . \tag{20}$$

First we prove Eq.(19). Note that by Eq.(13),

$$V[E[\chi/N_1N_2\,|\,(p_i),(\pi_i)]] = V[\Sigma_{i=1}^{n}p_i\pi_i] = E[(\Sigma_{i=1}^{n}p_i\pi_i)^2] - (E[\Sigma_{i=1}^{n}p_i\pi_i])^2.$$

The use of previous results gives

$$E[(\Sigma_{i=1}^{n}p_i\pi_i)^2] = E[\Sigma_{i=1}^{n}p_i^2\pi_i^2] + 2E[\Sigma_{i=1}^{n-1}\Sigma_{j=i+1}^{n}p_ip_j\pi_i\pi_j] = \frac{s_2^2}{n} + \frac{(1-s_2)^2}{n(n-1)}\ ,$$

$$E[\Sigma_{i=1}^{n}p_i\pi_i] = \frac{1}{n}\ .$$

Hence,

$$V[E[\chi/N_1N_2\,|\,(p_i),(\pi_i)]] = \frac{s_2^2}{n} + \frac{(1-s_2)^2}{n(n-1)} - \frac{1}{n^2} = \frac{(s_2 - \frac{1}{n})^2}{n-1}\ .$$

This proves Eq.(19). Eq.(20) follows therefrom by a comparison of Eqs.(17) and (18). Notice, the latter variance component goes to 0 as N_1 and N_2 go to infinity.

The error

Kullback's Eq.(21.7) is in error. The reason is an incorrect evaluation of the mean of the term $2\cdot\Sigma_{i=1}^{n-1}\Sigma_{j=i+1}^{n}p_ip_j\pi_i\pi_j$ in our Eq. (14). Kullback makes use of the identity

$$2\cdot\Sigma_{i=1}^{n-1}\Sigma_{j=i+1}^{n}p_ip_j\pi_i\pi_j = (\Sigma_{i=1}^{n}p_i\pi_i)^2 - \Sigma_{i=1}^{n}p_i^2\pi_i^2\ ,$$

but erroneously asserts that the mean of the right-hand side equals $1/n^2 - s_2^2/n$. True, $E[\Sigma_{i=1}^{n}p_i^2\pi_i^2] = s_2^2/n$. However,

$$E[(\Sigma_{i=1}^{n}p_i\pi_i)^2] \geq (E[\Sigma_{i=1}^{n}p_i\pi_i])^2 = 1/n^2,$$

with equality holding if and only if $p_i = \pi_i = 1/n$ for all i.

It follows that the mean of the left-hand side (2 x doublesum), and hence $E[\chi^2]$ and $V[\chi]$, have been underestimated. A simple calculation shows that the underestimation of $V[\chi/N_1N_2]$ amounts to

$$d = (1 - \frac{1}{N_1})(1 - \frac{1}{N_2}) \frac{(s_2 - \frac{1}{n})^2}{n-1}.$$

Thus, $d > 0$ except when $N_1 = 1$, or $N_2 = 1$, or $s_2 = 1/n$, that is, if $p_i = \pi_i = 1/n$ for all i. For large N_1 and N_2, d will nearly equal the variance of the conditional mean, given (p_i) and (π_i), see Eq.(19).

Finally, we note that the error through $V[\chi]$ has caused an error in the formula for the variance of $\Phi = \Sigma_{i=1}^{n}(a_i+b_i)(a_i+b_i-1)$, also in the case of non-matching distributions, see [1, p.49, Eq.(20.4)].

PARTIALLY MATCHED DISTRIBUTIONS

The two cases already discussed, namely matching and non-matching distributions, are the extreme cases in a general model in which any m, $0 \leq m \leq n$, pairs of identical probabilities are matched in advance, and other probabilities are paired randomly.

For the general model we shall give a formula for the mean of χ/N_1N_2. Let m denote the number of matches, and call the matched probabilities $h_1,\ldots,h_m$. That is, we assume $p_1 = \pi_1 = h_1,\ldots, p_m = \pi_m = h_m$, while the remaining p's and π's are randomly paired. As the first step we easily derive

$$E[\chi|(p_i),(\pi_i)] = N_1N_2(\Sigma_{i=1}^{m}h_i^2 + \Sigma_{i=m+1}^{n}p_i\pi_i).$$

Clearly, for $k > m$, $E[p_k] = E[\pi_k] = (1 - \Sigma_{i=1}^{m}h_i)/(n-m)$, so that

$$E[\chi/N_1N_2] = \Sigma_{i=1}^{m}h_i^2 + \frac{(1 - \Sigma_{i=1}^{m}h_i)^2}{n-m} \quad (0 \leq m \leq n-1) \tag{21}$$

For m=0, the sums are empty, and $E[\chi/N_1N_2] = 1/n$. For m=n-1, we have that $E[\chi/N_1N_2] = \Sigma_{i=1}^{n}h_i^2 = s_2$. In effect, this is the other extreme case, since n-1 matches imply n matches. Thus Eq.(21) is the general formula.

We shall demonstrate that increasing the set of matching probabilities never results in a lower value of $E[\chi/N_1N_2]$. It will suffice to show that the addition of any previously unmatched probability to an arbitrary set of matching probabilities will not lead to a lower $E[\chi/N_1N_2]$.

Let $E[\chi/N_1N_2|m]$ denote the mean of χ/N_1N_2 for m matched probabilities equal to $h_1,\ldots,h_m$, and let $E[\chi/N_1N_2|m+1]$ denote the mean of χ/N_1N_2 for m+1 matched probabilities equal to $h^1,\ldots,h^m,h^{m+1}$. Assume $0 \leq m \leq n-2$. Using Eq.(21) we find, after a little manipulation, the following simple formula:

$$E[\chi/N_1N_2|m+1] - E[\chi/N_1N_2|m] = \frac{n-m}{n-m-1}\left(h_{m+1} - \frac{1 - \Sigma_{i=1}^{m}h_i}{n-m}\right)^2. \qquad (22)$$

Hence, the difference is nonnegative as asserted.

REFERENCES

1. Kullback, S. 1976. Statistical Methods in Cryptanalysis. Laguna Hills, CA: Aegean Park Press. (Reprint, original in 1938).

2. Rao, C. R. 1973. Linear Statistical Inference and Its Applications. 2nd ed. New York: John Wiley and Sons.

SOFTWARE PROTECTION FOR MICROCOMPUTERS

JOHN M CARROLL AND PIERRE G. LAURIN

ABSTRACT: The most popular microcomputers have serious security weaknesses. Their password protection gives no protection at all against anyone skilled in the art of systems programming. The system described relies upon file encryption. It is implemented entirely in software and affords moderate security without incurring high overhead or memory residence costs.

INTRODUCTION

The migration of the microcomputer into the business office has forced a reexamination of the fundamental premises of computer security. (We understand the term microprocessor to refer to a central processing unit realized on a single semiconductor chip; a microcomputer consists of one or more microprocessors and the memory and gate circuits associated with it or them.)

Most computer security regimes rely on the principle of forced collusion, that is, that it is much harder to subvert two or more people than one. Thus we decree that programmers must not operate; operators must not program; tape librarians must assume custody of magnetic media removed from the main frame; and at least two systems programmers must sign off on security significant changes to software.

A second safeguard is multi-state operation of the computer. The computer is able to recognize at least two states, sometimes called supervisor state and problem state, but also called privileged/non-privileged, executive/application, or monitor/user. Security significant actions such as input, output, memory management, and access to system resources are handled in supervisor state by control programs functioning in accordance with specified access rules.

The principles of forced collusion and multi-state operation overlap in that multi-state operation places the control programs beyond the control of the user and that these control programs are implemented and maintained by trusted systems programmers.

In many business applications of microcomputers neither of these safeguards exists. Many popular micros do not support two-state operation, and a single individual may be responsible for both operating and programming.

OBJECTIVES

The objective of our research is to determine whether security controls can be imposed on an 8-bit microcomputer. The controls we developed act within the following scenario:

-There are assumed to be three classes of users:

(1) the trusted user (probably only one person, the boss, who possesses access rights to all information);

(2) the semi-trusted users who have been delegated the right to read specified items of sensitive information;

(3) the untrusted users who may not access any sensitive information.

-The security system to be described is a file encryption system. The information to be protected resides on five-and-a-quarter inch diskettes. The information is encrypted and is protected by three passwords: a system password, a disk password, and 350 record passwords. The name of the system is SCRAMBLE.

-The trusted user encrypts the protected files. The trusted user alone possesses the system password and in general has sole possession of the disk passwords, one for each encrypted diskette.

-The trusted user must open encrypted diskettes after which semi-trusted users can read them.

-Untrusted users can use the computer for program development and processing of non-sensitive information. The security provisions of the computer's operating system (disk password, access and update passwords, and provision for invisible files) are available to them as well as to the trusted and semi-trusted users.

CONVENTIONAL PASSWORD PROTECTION

This work was carried out using a Tandy Radio Shack TRS-80 Model I, Level 2 microcomputer operating under TRS-DOS 2.1 or 2.3 (Disk Operating System) with 48,000 bytes of random-access memory and two diskette drives.

Here is how conventional password protection works on the TRS-80:

-Each disk has a master password. This password can override any individual file password except for invisible and system file passwords. The master password can be assigned to all user files. It then overwrites the existing passwords, thus removing them.

-Each file can be assigned two separate passwords. These are the access and update passwords. The access password limits access as specified by a sublevel. The highest sublevel, kill, grants total control. The remaining sublevels are: rename, write, read, and execute. Any sublevel grants the functions of the sublevels below it. Passwords are hashed before storing on the disk.

-The invisible file feature creates a subdirectory in which the user can hide files. The directory command will not list invisible files.

-TRS-DOS protects its own file under a system file flag. When the permanent part of TRS-DOS calls a system file in response to an O/S (Operating System) command, the file will not execute unless the system's flag is set.

-The disk directory is composed of ten sectors (two "granules"). Each sector is read protected. TRS-DOS does not require read protection on system files.

SECURITY WEAKNESSES OF MICROCOMPUTERS

The principal security weakness of 8-bit microcomputers is that the computer cannot distinguish between operating-system instructions and user-program instructions. Among other things, this implies that the computer's entire memory is accessible by the programmer.

A closely related weakness is the absence of a system's stack. Both the system and the user use the same stack to store the return address from subroutine calls, to preserve the contents of registers, or to pass arguments to subroutines.

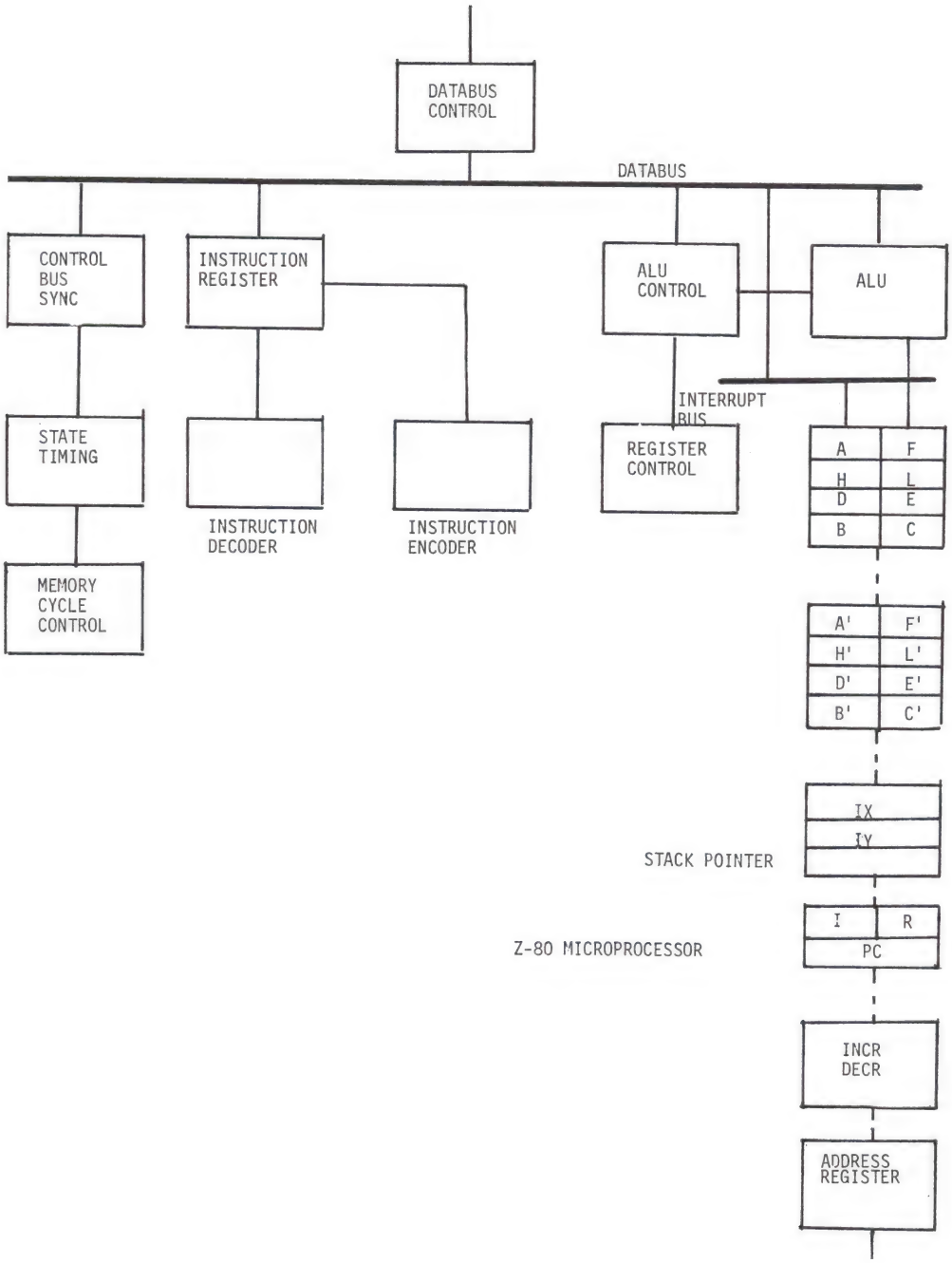


Figure 1. Simplified block diagram of Z-80 microprocessor.

The stack pointer (see Figure 1) is a 16-bit register that contains the address of the next available location in an external (i.e., exists somewhere in random access memory, not in the central processing unit) push-down/pop-up stack (i.e., an array of consecutive memory locations in which data can be stored in a first-in, first-out fashion). The stack can have any address (i.e., location) that is convenient. The stack can be used to store and retrieve the contents of the accumulator (shown in our diagram as part of the arithmetic logic unit), flags, program counter and any of the eight general-purpose registers (i.e., AF to B'C'). The stack pointer controls the addressing of the stack. The stack permits the processor to handle multiple-level interrupts and subroutine nesting. It stores the current state of the processor so that control can be returned after an interrupt or subroutine call.

Sixteen-bit microprocessors like the Motorola 68000 have two stack pointers and two stacks. One stack pointer and stack is dedicated to use by the operating system and is normally not under the programmer's control.

In a single-stack processor like the Z-80 (i.e., the "guts" of the Radio Shack TRS-80) a programmer can modify the execution of a system's command by changing the contents of the stack. Consider, for example, the handling of an interrupt. The programmer can push the address of his program down the stack, then create an interrupt. This forces the system to jump to the subroutine that handles it. But since the return address has been changed, control passes to the user's program.

Each device is assigned a port and an I/O (Input/Output) controller; the addresses are in ROM (Read-Only Memory). Inasmuch as the programmer has access to the entire memory including ROM, he can bypass the O/S and read or write to any I/O device thus circumventing any software protection. A programmer who learns his disk controller code can read or write any sector on any disk track. Thus he can get an entire file even if it is protected by a password on the directory.

SPECIFIC ATTACKS ON MICROCOMPUTERS

Special programs widely circulated in the microcomputer underground can help a would-be intruder.

SUPERZAP can copy any disk sector into memory where the intruder can modify it. The modified sector overlays the original when recopied onto the disk. When a read-protect sector is modified, the write routine must include the read-protect feature. SUPERZAP permits modification of read-protected sectors without disabling the read-protect flag as well as enabling the flag on any sector. This feature is useful when an intruder wants to alter a directory.

RSM 2 can disassemble a compiled program. It can read/write to disk but it can neither keep nor enable the read-protect flag. With RSM-2, a program can be loaded from any disk sector, disassembled and then modified by changing its object code on the disk.

One way to defeat a system protected by hashed passwords is to create a file with no password (default is eight blanks). This hashes to 9642 9642 hexadecimal on the TRS-80. Using SUPERZAP, the intruder can enter the target directory, overwrite any individual password with 9642 9642 hexadecimal and the corresponding file becomes unprotected. However, if the user is clever, he will cover up his theft of data by replacing the hashed version of the original password.

The Trojan Horse program can be a powerful tool for an intruder. It is an ordinary looking program that does what it is intended to do and does something to help the intruder as well. Suppose a program exists that is intended to update a sensitive file. Naturally it must be able to read from or write to it. If an intruder can get to that program, he can modify it so that in addition to updating the sensitive file, it also produces a copy for the intruder and stores it on a file in a surreptitious repository that nobody but the intruder knows.

OVERVIEW OF SCRAMBLE

SCRAMBLE is a storage tool. The user is not allowed to execute programs under SCRAMBLE supervision.

SCRAMBLE cannot be used to protect "execute-only" disks.

It is a password system and the passwords must protected. Passwords should be selected so they are not easily guessed or subject to discovery by exhaustive trial of permutations. Passwords should be eight characters long.

No user can change the system password. It is a permanent feature, unique to each copy of the SCRAMBLE system.

SCRAMBLE can be described as a new operating system with enhanced protective features. It results from a modification of existing TRS-DOS I/O routines to permit ciphering and deciphering. It retains all file manipulation commands, and disables certain commands that are undesirable in a secure environment.

We will describe SCRAMBLE on the operational, functional, and design levels--what it does, how it does it, and how it works.

OPERATIONAL DESCRIPTION

TO CREATE A PROTECTED FILE:

NOTE

INFO/TXT is the clear text version of a sensitive file

INFO/CRY is the enciphered version of that file

TRS-DOS is the normal operating system

1. Store clear text. With TRS-DOS and blank on drive 0, write INFO/TXT on drive 0.

2. Load SCRAMBLE. TRS-DOS+INFO/TXT on drive 0, and SCRAMBLE on drive 1.

3. Encrypt INFO/TXT. With TRS-DOS+INFO/TXT on drive 0 and a blank disk on drive 1, run FORMAT, create the disk password; COPY INFO/CRY on drive 1 and create a TRS-DOS file password.

4. Kill INFO/TXT.

(Note: The plus sign denotes concatenation).

TO READ A PROTECTED FILE

1. Load SCRAMBLE. TRS-DOS+blank on drive 0, and SCRAMBLE on drive 1.

2. Decrypt INFO/CRY. With TRS-DOS+blank on drive 0, and INFO/CRY on drive 1, run CHANGE giving the disk password; COPY INFO/TXT on drive 0 giving the TRS-DOS file password.

3. Read INFO/TXT.

4. Kill INFO/TXT.

ENCRYPTION

drive 0 = TRS-DOS 2.1 or 2.3 drive 1 = SCRAMBLE
INFO/TXT DES/SIM
SCRAMBLE/CMD

```
Command:        DES/SIM
System asks:    SYSTEM PASSWORD?
You answer:     XXXXXXXX (You get four chances with the prompt
                         then return to TRS-DOS with the
                         message NOTHING DONE
System:         DOS READY
                Put a blank disk on drive 1
Command:        FORMAT
System asks:    WHICH DRIVE?
You answer:     1
System asks:    NAME?
You answer:     TRS
System asks:    DATE?
You answer:     MM/DD/YY
System asks:    DISK PASSWORD?
You answer:     YYYYYYYY
System asks:    DO YOU WANT ANY TRACKS LOCKED OUT?
You answer:     N
System:         FORMATTING COMPLETE (file is "invisible")
Command:        CHANGE
System asks:    DISK PASSWORD?
You answer:     YYYYYYYY
System:         DOS READY
Command:        COPY INFO/TXT:0 TO INFO/CRY:1
Command:        KILL INFO/TXT:0
Command         ATTRIB INFO/CRY:1 (I,
                                   ACC = ZZZZZZZZ
                                   UPD = WWWWWWWW,PROT=level)
```

DECRYPTION

```
                drive 0 TRS-DOS      drive 1 = SCRAMBLE
Command:            DES/SIM
System asks:        SYSTEM PASSWORD?
You answer:         XXXXXXXX
                    Put the ciphered disk on drive 1
Command:            CHANGE
System asks:        DISK PASSWORD?
You answer:         YYYYYYYY
System              DOS READY
Command:            COPY INFO/CRY:1 TO INFO/TXT.ZZZZZZZZ:0
-read INFO/TXT-
Command:            KILL INFO/TXT,ZZZZZZZZ:0
```

FUNCTIONAL DESCRIPTION

Figure 2 shows that the system password is the first line of defense. It is a permanent feature unique to each copy of SCRAMBLE and should remain in the exclusive possession of the trusted user.

Unique disk passwords are assigned by the trusted user to individual diskettes as they are formatted. Some disk passwords may be delegated to semi-trusted users.

The access and update (with protection level) passwords and the invisible file feature are normal TRS-DOS protection mechanisms. They are assigned to the deciphered copy of the encrypted disk and are delegated to semi-trusted users.

The ciphering algorithm is transparent to the user. It consists of adding (modulo two) a 256-byte random key to each record. The whole buffer is ciphered (or deciphered) at each I/O call.

The same key is used for each of the 350 records on a disk but the key is shifted for each record by the output of a sensitive file shifting algorithm (a random-number generator) whose input is obtained by hashing the disk password and the record number.

The transposition matrix is used to prevent an intruder from knowing where to put information. For example, file invisibility depends on the setting of a single bit. If the intruder knows its location, he can reset it despite encryption of the file. The transposition matrix is a permanent feature unique to each copy of SCRAMBLE. The ciphering algorithm requires both

ciphering and deciphering matrices but one of these is generated from the other as it is needed.

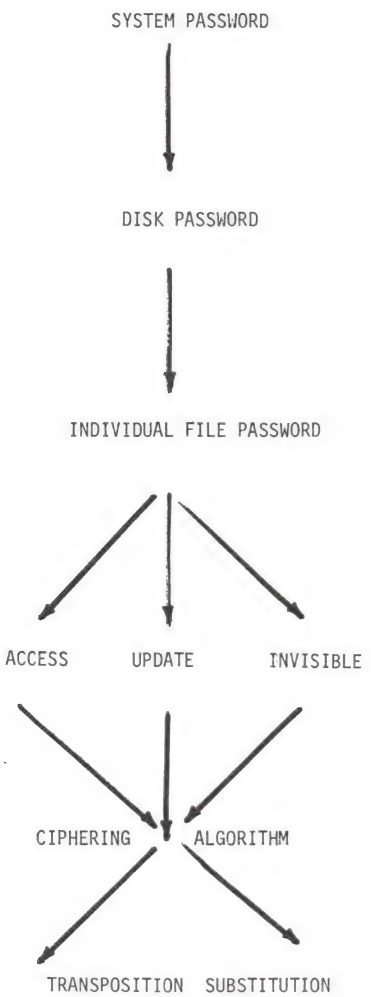


Figure 2. Protective mechanisms in order of application.

Figure 3 and 4 comprise a flow-chart of the process. Figure 5 symbolically illustrates the storage layout of system components DES/SIM, SCRAMBLE, and modified TRS-DOS. SCRAMBLE is delivered and stored in encrypted form. It is encrypted in the U.S. National Bureau of Standards Data Encryption Standard. A software implementation of DES is stored in DES/SIM.

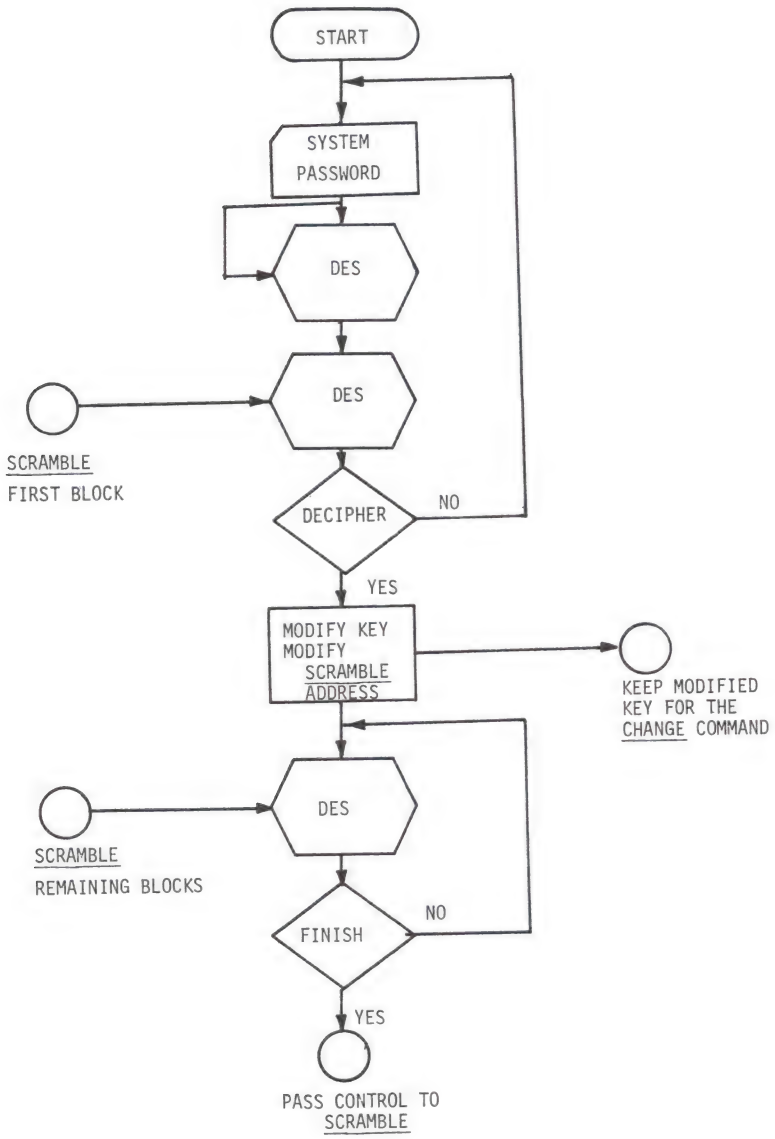


Figure 3. Flow chart, Part I.

The system password is used simultaneously as key (DES KEY 1) and input to DES. The output of DES is fed back in and used as the key (DES KEY 2) for the next step. This step consists of deciphering the first block of SCRAMBLE.

The first block of SCRAMBLE is really the first block of DES. It contains a routine to modify the DES key (thus creating DES KEY 3). It also contains the real starting address of SCRAMBLE). The address of SCRAMBLE stored in the clear in DES/SIM is fictional. It is intended to mislead an intruder.

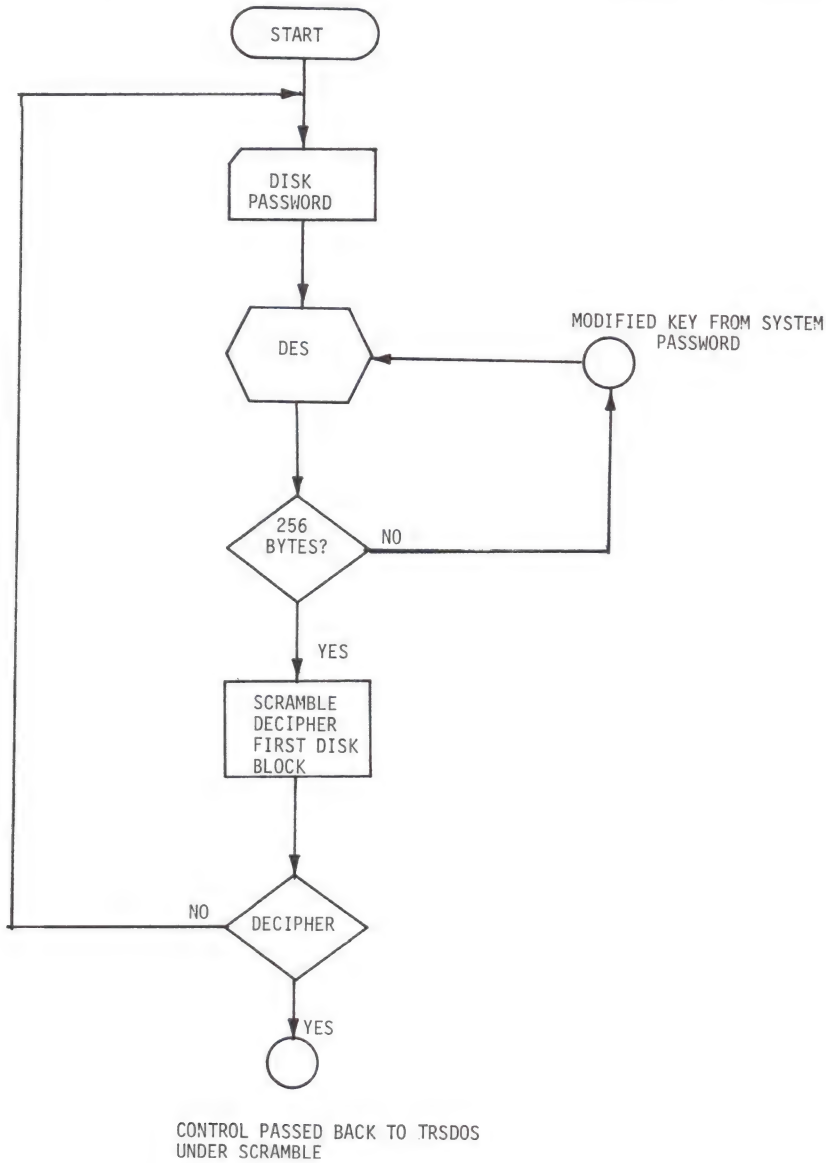


Figure 4. Flow chart, Part 2.

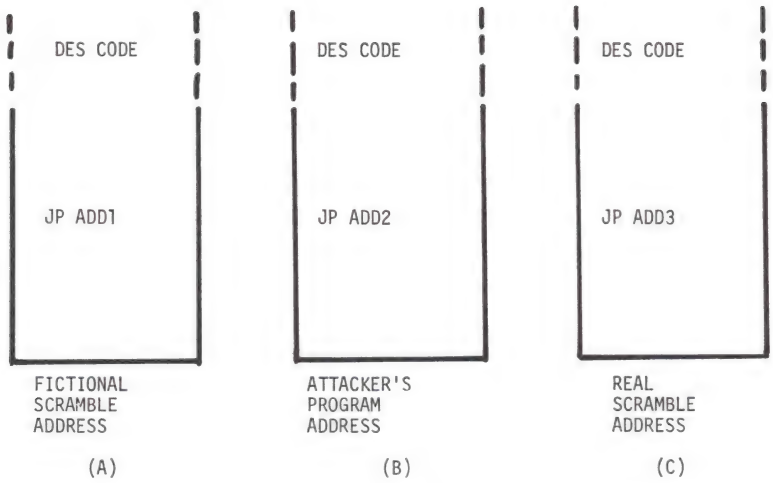


Figure 5. Protection of the address of SCRAMBLE:
(a) Fictional SCRAMBLE address,
(b) Attacker's program address,
(c) Real SCRAMBLE address.

Figure 6 shows how this deception works. A logical attack on SCRAMBLE would include changing the address of SCRAMBLE (as it resides in the clear text part of DES) to that of the intruder's program. However, in the loading process, the fictional starting address is overwritten with the real starting address (from block one of SCRAMBLE); an intruder's misdirecting address would be overwritten too, thereby avoiding an undesired transfer of control.

The reason we modify DES KEY 2 is to foil an intruder who implants his own copy of DES so as to bypass control transfer protection (i.e., the system password). If the intruder thereby succeeds in recovering DES KEY 2 and is unaware of the key modification, he will be forestalled from further penetration.

The "test for TRS-DOS disable flags" noted in Figure 5 tests to see that TRS-DOS commands TRACE or DEBUG are not enabled. If they are, the procedure will terminate. This safeguard is included to prevent an intruder from using these system utilities to circumvent our protective mechanisms.

Returning to Figure 3, we see that DES KEY 3 is inserted; DES proceeds to SCRAMBLE's real starting address, and deciphers the program.

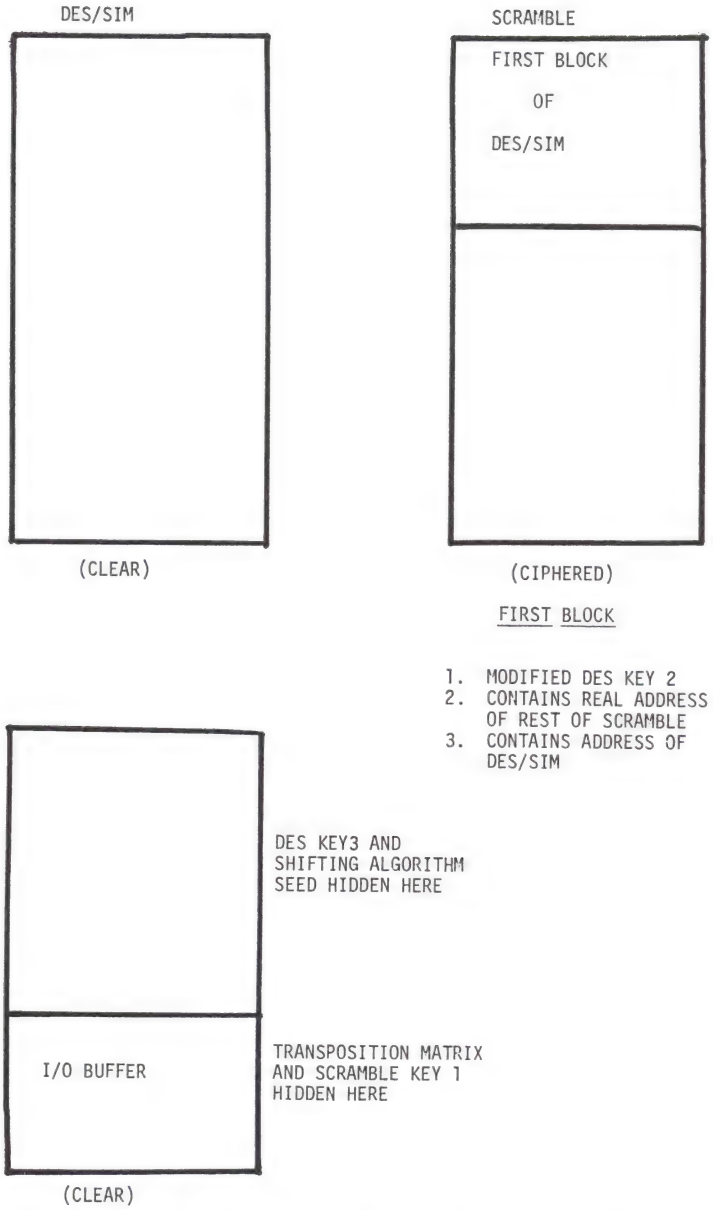


Figure 6. Storage of security software components.

Control is now passed to SCRAMBLE. It first modifies TRS-DOS by disabling TRACE, DEBUG, LOAD and BASIC. It supplies three new O/S commands; FORMAT, CHANGE and EDIT, overwriting TRS-DOS DATE and TIME to make room for them.

With DES KEY 3 inserted in DES, the disk password is now repeatedly input to produce a 256-byte key (SCRAMBLE KEY 1).

FORMAT allows the user to format a blank disk which will receive encrypted information. When formatting a disk, SCRAMBLE will encipher BOOT/SYS and the directory file by adding them modulo 2 to SCRAMBLE KEY 1.

CHANGE permits the trusted user to work on different ciphered disks without reloading SCRAMBLE. It asks for the disk password and creates SCRAMBLE KEY 1 just like FORMAT does.

EDIT is the Z-80 editor. It permits a trusted or semi-trusted user to make alterations on an encrypted file without having to remove it from the protected disk.

DESIGN DESCRIPTION

SCRAMBLE uses SCRAMBLE KEY 1 to encrypt and decrypt text. The shifting algorithm is used to create 256 different record keys. (We can regard them as SCRAMBLE KEYS 3 to 257). The shifting algorithm is a pseudo-random number generator. It is given a unique seed for each record and produces a 256-byte pseudo-random number sequence which is added modulo 2 to the record. To make the sequences different for every ciphered disk, the disk password is hashed with the record numbers to create the seeds. The record number is the track number multiplied by ten plus the sector number.

There are 1.37 records per key (350/256). The keys are called randomly. The seeds consist of 3-bytes. Byte 1 is the second byte of the hashed disk password. Bytes 2 and 3 are the sum of the hashed disk password and the record number.

Assume the disk password is SCRAMBLE. This is hashed by an algorithm that reduces the eight bytes to two. SCRAMBLE hashes to A21A hexadecimal. The first byte of the seed is 1A. Assume we are accessing the record on track ten, section five. The record number is 0069. Bytes two and three are A21A+ 0069 = A283. The seed is, therefore, 1AA283.

Figure 4 shows that after the disk blocks have been deciphered (enciphered), control passes to TRS-DOS.

Figure 6 shows that the transposition matrix and SCRAMBLE KEY 1 are hidden in the O/S I/O buffers. They are continually being destroyed and rebuilt. Furthermore, if an intruder should stop the execution of SCRAMBLE, neither component will be rebuilt. Before any I/O, SCRAMBLE takes the key and matrix out

of the buffer. At the end of the routine, they are recopied into the buffer. Every O/S operation must perform at least one I/O routine; so if an intruder disables SCRAMBLE, the loading of the O/S will destroy both the key and the matrix.

DES KEY 3 is stored in eight bytes used by the O/S. The seed of the shifting algorithm is stored in two locations used by the O/S. Thus they are afforded the same protection accorded SCRAMBLE KEY 1 and the transposition matrix.

After the text is decrypted, control is returned to (modified) TRS-DOS. The computer can be entrusted to the semi-trusted user who can execute any TRS-DOS command except the four disabled ones. When the semi-trusted user completes his session, he has to reset the computer thus destroying DES KEY 3, SCRAMBLE KEY 1, the seed for the pseudo-random number generator, and the transposition matrix.

RESULTS

SCRAMBLE requires 29 disk sectors or 7K of memory.

The time overhead added to the system by the measures taken to protect the crypto parameters is 500 microseconds per I/O call. The total overhead per I/O call is about 8.5 milliseconds. A normal I/O routine takes on the average 400 milliseconds including displacement of the disk read/write head. Thus the ciphering algorithm adds only 2.15 percent overhead to each I/O call.

VULNERABILITIES

There are two known vulnerabilities to SCRAMBLE.

1. An intruder can overwrite a file with junk. The only effective protection against this kind of attack is physical protection of the disks.

2. A Trojan Horse program can be used. In such an attack TRS-DOS can be modified to bypass SCRAMBLE.

Here's how it could work. Using RSM 2 the intruder can examine SYS1/SYS (the permanent part of TRS-DOS) and discover the starting address of some appropriate command. Next, he could replace that TRS-DOS subroutine with a copy of RSM-2. Then he can let the trusted user load SCRAMBLE and create the necessary crypto keys. Now if the intruder is one of the semi-trusted users, he can call the TRS-DOS command he has modified and examine SCRAMBLE until his heart's content.

CONCLUSIONS

Like every protective mechanism, SCRAMBLE can be defeated by an intruder who is dedicated, resourceful, and enjoys a fair amount of good luck.

Within its limitations, SCRAMBLE is useful for controlling the dissemination of sensitive information. It enables the trusted user to decide when and to whom information residing on encrypted disks shall be released and does so without the need for expensive hardware and with extremely low overhead.

We believe our development procedure that consisted of repeated cycles of measure-attack-countermeasure should be of interest to designers of software security systems that are not necessarily directed against the same threat scenario.

Some of the newer 16-bit microprocessors provide for two-state operation but they have yet to be incorporated into popular microcomputers. Even when, and if, they are, file encryption would still be valuable as a back-up protective mechanism. Only a multi-state machine with built-in encryption/decryption chips would obviate the need for systems like this.

[The research reported in this paper was carried out with the support of the Natural Sciences and Engineering Research Council of Canada under grant No. A7132 and the Canadian Certified General Accountants' Association under grant No. CG4501 whom the authors gratefully thank.]

REFERENCES

1. Barden, Jr. W. 1979. The Z80 Micro-Processor Hardware. Indianapolis: Howard W. Sams and Co, Inc.

2. Canning, R. G., and B. McNulling. 1978. Micros invade business world. Datamation. August.

3. Carroll, J. M. 1977. Computer Security. Woburn, MA: Butterworth Publishers, Inc.

4. Cassell, D. 1978. Putting micros into perspective. Datamation. August.

5. Clarke, W. 1975. Value of access control. Data Processing. Jan - Feb.

6. Davis H. A. 1979. Comparative architecture of three 16-bit microprocessers. Computer Design. July.

7. Dennis, C. 1973. Security vs. performance. Datamation. November.

8. Dollhoft, T. 1979. 16-bit Micro-Processor Architecture. Reston, VA: Reston Publishing Co.

9. FIPS No. 46. 1976. Washington: National Bureau of Standards.

10. Hoffman, L. 1977. Modern Methods for Computer Security and Privacy. Englewood Cliffs, NJ: Prentice-Hall, Inc.

11. Gait, J. 1978. Easy entry: the password encryption problem. Operating System Review. 12: 3.

12. Konheim, A. G., M. H. Mack, R. K. McNeil, B. Tuckerman, and G. Waldbaum. 1980. The IPS cryptographic programs. IBM Systems Journal. 19: 2.

13. Laurin, P. G. 1980. Micro-Computers: How Safe? Master's Thesis. The University of Western Ontario. London, Ontario.

14. Mennie, D. 1978. Personal computer for the entrepreneur. IEEE, SPECTRUM. September.

15. MICROSOFT ADVENTURE GAME. Microsoft Consumer Products. Bellevue, WA.

16. Pennington, H. C. 1979. The TRS-80 disc and other mysteries. Upland CA: LJG Inc.

17. Popek, G., and C. Kline. 1979. Encryption and secure computer network. Computer Surveys. December.

18. Simmons, G. J. 1979. Symmetric and asymmetric encryption. Computing Surveys. December.

19. Sugarman, R. 1979. Computer: Our micro universe expands. IEEE, SPECTRUM. January.

20. Sugarman, R. 1979. On foiling computer crime. IEEE, SPECTRUM. July.

21. Sykes, D. 1976. Protection of data by encryption. Datamation. August.

22. Van Tassel, D. 1969. Advanced cryptographic techniques for computers. Communications of the ACM. December.

23. Williams, D. 1974. Cypher techniques keep files confidential. Canadian Datasystmes. 6: 8.

24. TRS-DOS and Disk BASIC Reference Manual. Fort Worth, TX 1979. Radio-Shack.

"It says 'Now you know how King Kong felt.'"

CORRECTIONS FOR PUBLISHED COPY OF UNITED STATES CRYPTOGRAPHIC PATENTS: 1861 - 1981

JACK LEVINE

[Ed. note: Cryptologia published United States Cryptographic Patents: 1861 - 1981 in 1983. Authored by Professor Jack Levine, Professor Emeritus of Mathematics, North Carolina State University, the book has served as the definitive work in this area. This book is still available from Cryptologia for $10.00. As with all such endeavours there will be errors in manuscript preparation despite all good intentions. Below is a list of the corrections which have been found for insertion in your copy of the book.]

Page	Patent	Correction
1	48,681	Change "Edward" in name to "Edwin"
1	294,175	Change "Cryptographical" in title to "Cryptographal"
4	797,016	Change "Pimental" in name to "Pimentel"
9	1,472,218	Change "receiver" in title to "receiving"
14	1,945,014	Change "20" in date to "30"
15	2,093,397	Change "4" in date to "14"
17	2,382,251	Change middle initial "E" in name to "D"
19	2,396,288	Change "V" in name to "v" to read "van"
22	2,479,338	Change "communication" in title to "communications"
24	2,586,475	Change "Vladimir" in name to "Wladimir"
26	2,689,686	In title, change "digraphs and trigraphs" to read "digraphs, trigraphs"
27	2,777,897	In name, change "Bretener" to read "Gretener"
29	2,816,156	In name, change "Fawley" to read "Pawley"
29	2,832,826	In date, change "20" to "29"
31	2,939,916	In title, change "translation" to read "translating"
32	3,000,486	Omit "et al."
33	3,033,922	In name, change "Oran" to read "Oren"
34	3,170,033	In title, change "signals" to read "symbols"
34	3,175,033	In date, change "25" to "23"
34	3,188,391	In name, change "Francis" to read "Francois"
36	3,234,663	In date, change "16" to "15"
36	3,309,694	Add "et al" at end of name

39	3,445,591	In name, change "Keohler" to read "Koehler"
39	3,490,044	In title, change "Communication" to read "Communications"
39	3,499,992	In title, change "communication" to read "communications"
47	3,953,677	In title, change "multiplex" to read "multiple"
48	3,980,836	In name, add "et al."
52	4,156,108	In title, change "transmission" to read "transmissions"
52	4,163,872	In title, change "signal" to read "signaling"
53	4,170,757	In title, change "Method and" to read "Method of and"
53	4,171,513	In title, change "communication" to read "communications"
54	4,179,658	In name, change "Blitzer" to read "Bitzer"
58	In Patent following 4,268,860, change "4,721,482" to read "4,271,482"	
63	Under "Druz", change "7,755,333" to read "2,755,333"	
66	Under "Gannett", change "2,983,326" to read "3,983,326"	
66	Under "Mathes", change "3,401,877" to read "2,401,877"	
68	Under "Atalla", change "2,283,599" to read "4,283,599"	

In addition there is one known omission:
4,156,314 Leo Rosen May 29, 1979 Rotors for a ciphering machine.

The following patent numbers are to be added to the list of Secondary Patents:

3,699,496 3,700,900 3,778,128 4,070,091 4,095,192 4,174,149.

Naval Intelligence

THE SLIDEX RT CODE

LOUIS KRUH

The Slidex Radiotelephone (RT) Code was one of almost two dozen codes and ciphers authorized by the Signal Division, Supreme Headquarters, Allied Expeditionary Force (SHAEF) for use in Operation "Overlord," the invasion of Normandy in 1944, by the combined US-British forces. [1]

The Slidex unit consists of a metal frame in which a card with 12 columns and 17 rows is placed. The card is preprinted with words, letters and numbers in each of its 204 boxes. Different cards are provided for different units, each with its own vocabulary, appropriate to the user.

The Slidex Frame

Across the top and down the side of the frame are slots for key alphabets. These provide the letters for bigram coordinates to encipher the words, letters (for spelling other words), and numbers on the card.

The Slidex was the only system authorized for use in radiotelephone conversations by airborne troops. Instructions issued February 1, 1944, specified that the keys "must either be based on a memorisable code word or be distributed in a manner that ensures secure encoding even though a proportion of such keys may have been captured by the enemy during the operation." [2]

The instruction booklet for the Slidex directed the user only to encode those portions of conversation which might be of value to the enemy. It specifically said that it was not to be used to encode the entire message. (See complete instructions following.) This proved to be its main weakness as the Slidex was broken soon after the Germans first intercepted traffic during maneuvers in southern England in March, 1944. [1] Subsequently, U.S. signal intelligence authorities declared that the Slidex "was cryptographically insecure in that it involved a mixture of code and clear text and was therefore particularly susceptible to cryptanalysis." Instructions were then issued to recommend a new method of indicating the key setting. The method was adopted on December 23, 1944, throughout the AEF but two weeks later, on January 6, 1945, the Signal Division recommended the Slidex be replaced within the U.S. forces. [3]

It was a relatively short-lived history for an undistinguished code made even more vulnerable by inept instructions.

REFERENCES

1. Thompson, G.R. and D.R. Harris. 1966. The Signal Corps: The Outcome. Washington: USGPO. pp. 90-91.

2. Supreme Headquarters, Allied Expeditionary Force. 1945. Report of Signal Division, SHAEF in Operation "Overlord." 4: 1175.

3. Supreme Headquarters, Allied Expeditionary Force. 1946. Report of Signal Division, SHAEF in Operation "Overlord."5: 1587.

[On the following pages we reproduce the Instructions for the Use of Slidex RT Code - text on pages 1-3, and 6-8, with the Imaginary Vocabulary List occupying pages 4-5 of the original Instructions manual.]

RESTRICTED

The information given in this document is not to be communicated, either directly or indirectly, to the Press or to any person not authorized to receive it.

INSTRUCTIONS FOR THE USE OF SLIDEX RT CODE

1. General

(*a*) This code will be used exclusively to conceal *those portions* of a RT or key conversation which it is considered might be of value to the enemy. It will NOT be used to encode the *whole* of a conversation unnecessarily.

(*b*) All officers and such other ranks as may have to carry on, transmit, receive or handle either type of conversation must know how to use the code.

2. Equipment

The equipment consists of a folding case, code cards and cursors, long and short. The case has a pocket in which the cards and cursors are kept. In use, the case is opened so that the pocket lies to the left. On the right is a frame consisting of two vertical metal strips to hold the card in use and two channels, one across the top and one down the left-hand side, into which the cursors slide.

3. Code Cards

(*a*) All code cards have 12 columns and 17 rows forming 204 rectangles. Separate cards are provided as under:—

Ops/Sigs, Med, RA, RE, RAC/REME, Air, Q(*a*) for use down to Corps rear links, Q(*b*) for use forward of Corps, Unit.

(*b*) Each type of card has,

(i) A vocabulary, printed in black, appropriate to the user, except the Unit card. The words, phrases, etc., are arranged alphabetically by rows. Unit cards are blank and it is the intention that each unit shall prepare and insert on its Unit cards a vocabulary suitable for use within its own unit.

(ii) Numbers 0—9, 00—99 and the letters of the alphabet printed in red in the top left-hand corners of the rectangles. The numbers 00—99 are arranged in numerical order by columns, interspersed with letters and single figures. The single figures are arranged so that 0 precedes 00, 1 appears between 09 and 10, 2 between 19 and 20, and so on. The letters are in alphabetical order, a complete alphabet to every four columns, thus each letter appears three times on the card. In one alphabet on each card E and T are duplicated and so appear four times each.

(iii) Twelve switches, six SWITCH ON and six SWITCH OFF, also printed in red.

4. Cursors

(*a*) Cursors are of two kinds,

(i) LONG, having divisions the width of a card column. They slide into the horizontal channel above the card and " horizontal keys " (q.v.) are written on them.

(ii) SHORT, having divisions the width of a row on the card. They slide into the vertical channel to the left of the card and " vertical keys " (q.v.) are written on them.

(*b*) Cursors, both LONG and SHORT, are known as BLACK, GREEN or RED according to the colour of their division lines. They are also marked at either end, on one side with a band of colour and on the other side with coloured square dots.

(*c*) LONG cursors have 16 divisions, SHORT cursors have 21 divisions.

5. Keys

The device may be used with either sliding keys or fixed keys, the latter being used with the UNIT card only.

(*a*) When sliding keys are used,

(i) A horizontal key and a vertical key (known as a " key pair ") are required. Also, one of the rectangles on the card will be designated as the " key rectangle " by which the settings of the cursors for a conversation are indicated.

(ii) Each horizontal key will consist of the first 12 letters of the alphabet (A—L) in jumbled order, with the first four letters of the jumble repeated at the end so that, when it is written on the cursor, each division will contain a letter.

(iii) Each vertical key will consist of the first 17 letters of the alphabet (A—Q) in jumbled order, with the first four letters of the jumble repeated at the end so that, when it is written on the cursor, each division will contain a letter.

(iv) With each key pair a " key rectangle ", chosen at random, will be issued. It will be indicated by giving the red number or figure in the chosen rectangle. As each letter appears on the card more than once, it will be necessary when the rectangle chosen contains a letter to indicate which particular rectangle is meant by specifying, *e.g.*, 1st N, 2nd N, 3rd N (*i.e.*, the first, second or third N in sequence on the card).

Example:—

Horizontal Key } G C B F J A L E I D K H G C B F

Vertical Key } I C H B K A G J N F O Q M P D L E I C H B

Key Rectangle 2nd N.

(*b*) When fixed keys are used, *i.e.*, with UNIT card only,

(i) A " key pair " only is required consisting of the first 12 letters of the alphabet in jumbled order for the horizontal key and the first 17 letters of the alphabet in jumbled order for the vertical key.

(ii) The repeated letters and key rectangle mentioned in para. 5 (*a*) are NOT required.

(*c*) It is essential that the letters of all keys be arranged haphazardly and that the sequence of letters vary from key to key. " Key rectangles " must also be chosen at random.

6. Issue of Keys

(*a*) Each divisional headquarters will prepare and issue key pairs and key rectangles as in para. 5 (*a*) for use by all holders down to unit level within the division. These keys will be written on BLACK cursors.

(*b*) Army headquarters will prepare and issue key pairs and key rectangles as in para. 5 (*a*) for use

(i) by all holders behind division down to unit level,
(ii) between divisions,
(iii) between divisions and higher formations.

These keys will be written on RED cursors.

(*c*) The Army Commander may, at his discretion, order the use of the " Army " key pair and key rectangle by all holders of one or more types of card; *e.g.*, holders of the RA card may be instructed to use the " Army " keys for conversations at all levels throughout the Army. In the interests of security this arrangement should be used as seldom as possible.

(*d*) Units will work with fixed keys and will prepare key pairs as in para. 5 (*b*) for use within the unit. These keys will be used with Unit cards and will be written on GREEN cursors.

Note.—Indelible pencil must not be used for writing on cursors.

7. Key Changes

Normally, key pairs and key rectangles will be changed daily at midnight, but whenever the volume of traffic makes it desirable on security grounds, keys may be changed twice daily at the discretion of the Army Commander. For this purpose

(*a*) Two key pairs and, preferably, a key rectangle for each, will be provided daily.

(*b*) The keys to be used for the first part of the day will be written on the sides of cursors end-marked with coloured DOTS; those for use during the remainder of the day will be written on the sides end-marked with a coloured BAND.

(*c*) The time for the change during the day will be the same throughout an Army and should be chosen so that approximately equal weights of traffic are thrown on each key pair.

8. Lateral Communication

Lateral communication at and behind divisional headquarters within the same Army will be carried out on the Army keys. All other demands for such communication will be met by an *ad hoc* passing of keys.

IMAGINARY VOCABULARY

	I													
G	C	B	F	J	A	L	E	I	D	K	H	G	C	B	F
H		O (Fig) Able	F Advance	17 Air	U Aircraft	32 Answer	42 Area	M Armoured	T Arrange	66 Arrive	G Assume	82 Attack	90 Battery	
B		OO Battle	09 Block(s)	M Bomb(er)	23 Bombline	33 Boundary	E Bridge	5 Br gade	58 By	A Cable	74 Can move	N Cancel	U Capture	
K		SWITCH OFF	G Carrier	18 Casualty	V Centre Line	A Change	43 Close	50 Closed down	U Come under	67 Comd. of	H Comm n	83 Complete	91 Confer	
A		A Confirm	SWITCH ON	N Contact	W Corps	34 Counter	F Cross(ed)	N Cut off	V D R	68 D R L S	75 Day	O Delay	92 Direction	
G		01 Diversion	H Division	SWITCH OFF	24 Dump	East	G Echelon	O Empty	59 Enemy	69 Ether	I Field	84 First	93 French	
J		B From	1 (Fig) Front	O Gas	25 German	35 Green	44 Ground	51 Guns	W H Q	B Harass	J Harbour	85 Head	V Heavy	
N		C Hold	10 Hospital	P Hour(s)	SWITCH ON	36 Hundred	H Hygiene	52 Immediate	6 Infantry	C Inform	76 Killed	86 Kms	94 Landing	
F		02 Leave	11 Lift(s)	Q Line(s)	26 Link	SWITCH ON	45 Locate	P Location	60 M G	D Map	77 Medium	P Meet	W Mile(s)	
O		03 Mine(s)	12 Minus	19 Minute(s)	27 Move	B N T R	SWITCH OFF	53 Near	X Night	7 Normal	K North	87 Now	X Object	
Q		D Observe	I Occupy	R Open(s)	28 Operation	37 Ordnance	I Pack	SWITCH ON	61 Pass	70 Patrol	78 Petrol	Q Post	95 Pound	
M		04 Prepare	13 Priority	2 R A F	X R V	C Rate	46 Recce	Q Red	SWITCH ON	E Regt	L Reinforce	88 Repair	96 Report	
P		05 Reserve	J Restore	20 Return	29 Right	38 Road	U Round(s)	54 Route	62 Salvage	SWITCH OFF	79 Same	R Section	97 Self	
D		06 Send	K Shell(s)	S Ship	Y Signal	39 Smoke	47 South	R Sp Gp	63 Spare	E Stock	SWITCH OFF	89 Stop	Y Stores	
L		07 Strength	14 Sup Pt	21 Tac R	Z Take	D Tank	48 Task	55 Telephone	Y To	71 Today	S Tomorrow	S Ton(s)	98 Track	
E		Tractor	15 Traffic	22 Trailer	3 Train	4 Transmit	K Transport	S Type	64 Unit	F Until	80 Use	T Verey Light	Z Very	
I		E w.e.f.	L Wait	T Watch	30 Water	40 Weather	49 Whether	56 Which	Z When	72 Well	M Went	SWITCH OFF	99 West	
C		08 Will	16 With	T Workshop	31 Wound	41 Yards	L Yellow	57 Yet	65 You	73 Your	81 Zero	9 Zone	SWITCH ON	
H														
B														

CARD 17

Key pair and key rectangle as in para. 5(*a*). Setting co-ordinate DA

9. Unit Vocabularies

Each unit will prepare for internal use with and between sub-units a vocabulary which will be written in the blank rectangles of unit cards in a convenient order. These vocabularies will be amended or changed completely, as necessary.

10. Distribution

The Army commander will, according to operational requirements, lay down what distribution is to be given to,

(*a*) The keys issued at all levels.
(*b*) The vocabularies produced by units for their own use.

11. How to use the Code

(*a*) Select from the cards provided the one appropriate to the conversation contemplated.

(*b*) Place the card in the frame. To do this insert the edge of the card under one of the vertical metal strips. Slightly bend the card and slip its other edge under the other strip.

(*c*) Insert the cursors, bearing the appropriate key. Users who require to use one key pair only will find it convenient to keep the cursors permanently in the channels.

(*d*) When sliding keys are used (BLACK or RED cursors) move the cursors, the vertical up or down, the horizontal left or right into any position relative to the card but ensuring that every row and every column of the card has a key letter opposite to it. The variation of co-ordinates from message to message given by this arrangement adds materially to the security. When fixed keys are used it is only necessary to adjust the cursors so that the keys are in the correct positions relative to the card.

(*e*) Normally, because of the net employed, the receiver will know which card the originator is using. Should there be any variation, the originator will give the card number (printed in the bottom right hand corner of the card) before beginning the conversation.

(*f*) If sliding keys are used, the originator of the conversation will indicate to the receiver the position in which he has set his own cursors by giving the co-ordinates of the " key rectangle " and allowing the receiver a sufficient interval to set his cursors before continuing. With fixed keys this is not necessary.

(*g*) Each phrase, word, letter or number which has to be concealed will now be encoded by taking the letter co-ordinates of the rectangle in which it appears.

(*h*) The first letter of a co-ordinate will be taken from the horizontal key, the second letter from the vertical key.

(*i*) Users are advised for quick reference to make a note of the key rectangle on a convenient part of one of the cursors to which it pertains.

(*j*) The phonetic alphabet will always be used when giving co-ordinates.

12. **Spelling**

(*a*) Words which do not appear in the vocabulary and which, for security reasons, cannot be mentioned in clear will be encoded by means of the red letters as follows:—

Give the co-ordinates for one of the SWITCH ON rectangles, if necessary, and then give the co-ordinates for each letter of the word to be encoded, concluding the spelling with the co-ordinates for one of the SWITCH OFF rectangles if necessary. Alternatives are provided for all letters and switches. *Full use* will be made of these alternatives.

(*b*) If a letter is repeated in a word, each of the repeats will be taken from a *different* alphabet.

13. **Figures**

(*a*) When figures have to be encoded the red numbers will be used as follows:—

Give the co-ordinates for one of the SWITCH ON rectangles, if necessary before encoding the figures and, if necessary, the co-ordinates for one of the SWITCH OFF rectangles at the end of the figures.

(*b*) Numbers of more than two figures will be encoded two figures at a time, the odd figure, if any, being encoded last. (*See* examples, para. 15.)

14. **Security**

(*a*) Whenever it is possible to do so without confusing the decoder, the use of either or both switches will be avoided. *E.g.*, when spelling or figures ends a conversation SWITCH OFF will not be used. An encoded passage which can be nothing but spelling or figures to the receiver, needs neither switch.

(*b*) Care should be taken to frame conversations so that the portions given in clear afford as little clue as possible to the nature of the encoded portions.

(*c*) Unit vocabularies should be compiled with a view to keeping the necessity for switching to a minimum.

(*d*) Formations and units responsible for issuing keys will provide emergency keys for use if and when required.

(*e*) The loss or compromise of any key or list of keys will be reported immediately to the issuing authority who will take the necessary steps to restore security.

(*f*) If circumstances arise in which this code is in imminent danger of capture, all keys will be burnt as a first priority, after which all vocabulary cards and the instructions will be destroyed by the same means. N.B.—The cursors are inflammable.

15. **Examples**

These examples are founded on the diagram on pages 4 and 5 and the keys given in para. 5 (*a*).

(*a*) *Vocabulary*.

Cancel move—DA CB LO.
Report location harbour—DA BM DF GJ.
What is your centre line—DA What is your LK.

(*b*) *Spelling*

NORTHAMPTON—DA KM* CB AJ AQ KH IN FA DH CF CE DG AA CI*.

(*c*) *Figures.*

29 DA DQ* LP AG*
300 DA EF* LI FH IO*
2004 DA LN* AP FM HP*
71625 DA BC* HL KP DB GD*
035289 DA JA* FO DN CD FK*

Note.—Switch co-ordinates (marked with an asterisk in the above examples) will not be used unless they are necessary. All the possible switch groups have been used in (*b*) and (*c*).

(*d*) With the omission of the co-ordinates of the key rectangle, *i.e.*, DA, the above examples give the coded version which would be obtained from a UNIT card having the vocabulary given in the diagram and key pair,

Horizontal—F J A L E I D K H G C B
Vertical—H B K A G J N F O Q M P D L E I C

(B43/193) 96000 1/44 G.S.St. C.1338

BRITISH INTELLIGENCE - VOLUME II -- BOOK REVIEW

RALPH ERSKINE

British Intelligence in the Second World War: Its Influence on Strategy and Operations, by F.H. Hinsley, E.E. Thomas, C.F.G. Ransom and R.C. Knight. London: Her Majesty's Stationery Office, 1981. 15.95 pounds sterling. New York: Cambridge University Press. $39.50. Volume Two. 850 pp.

This volume starts in mid-1941, where Volume I, reviewed in Cryptologia (January, 1982), [1] left off. It continues the story until mid-1943, covering all areas except the Far East, which was excluded because it was so much within the United States' sphere of influence and inadequately included in British records. The war at sea and in the air are dealt with comprehensively. On land, the North African campaign naturally falls for the most attention.

For those who were disappointed by the style or even by some of the content of Volume I, this volume comes as a pleasant surprise. That volume made rather heavy reading. Overall, Volume II ranks among the better written of the British series, even if it still cannot be said to be lively or anecdotal. Official histories, being works of record, have to set out a multitude of facts, which does not lend itself to a light style.

When reviewing Volume I, William P. Bundy complained of the lack of references to individuals and of the fact that there was no bibliography [1]. The latter will, we are assured, be remedied in Volume III. The former is typical of most of the British official histories, many of which had the same general editor as this volume. And this reviewer has much sympathy with the view of the authors that such a work should follow Flaubert's precept: pas de monstres, et pas de héros. Given the subject matter, how could any general history distinguish between the contributions made by the many individuals at the Government Code and Cypher School ("GCCS" or "Bletchley") to the breaking of Enigma? A further reason for the emphasis on organizations is that the work has been based mainly on written records. We can hardly complain since, given the size of the task, it could scarcely have been otherwise. Other

writers can fill in the interstices by tracking down and interviewing those of the participants who are still alive.

Although the authors had virtually unrestricted access to official files, the constraints included a bar on publication of "details of the methods by which [intelligence] was obtained" [2]. There is therefore no description of the cryptanalysis of Enigma or Geheimschreiber (Fish) by GCCS. So we will have to content ourselves with accounts elsewhere [3].

Readers of Cryptologia will probably search first for the treatment of cryptographic topics. There is material in plenty. The index is a good guide to the work's coverage. The entry for GCCS covers almost 5 columns, that for Enigma, 7. Some items are short, and poignant, histories in themselves ("Italian Navy — GC and CS breaks cyphers of...; GC and CS loses cyphers of...;"). There are several columns on Sigint and a few inches on C38m (modified Hagelin).

However, the book is about much more than cryptology. Essentially, it deals with the organization of intelligence and its integration into the command structure. No single element of intelligence, no matter how important, could sensibly be considered and used on its own. Enigma-based intelligence ("Ultra" in British terms) was supplemented by lower-grade intelligence collected by the Y Services (nets of listening posts established by the Army and Air Force [4]), electronic intelligence [5], traffic analysis and other measures: all played their part. And then the results had to be analyzed, distributed and used. Perhaps that is why there is, and had to be, so much discussion of committees and structure in these volumes. It was in the efficient organization and use made of all intelligence that the Allies excelled.

The introductory four chapters (which start with chapter 15) consider developments in the organization of intelligence. Chapter 16 describes arrangements with the United States and relations with the Soviet Union. While there were some initial differences with the United States Navy Department, by June 1943 there was very full collaboration. A visit to the United States by GCCS in October 1942 ensured that the Navy Department would receive Enigma decrypts and technical assistance from the British. The Navy Department agreed to supply GCCS with Japanese Naval decrypts and other intelligence and allowed GCCS to coordinate work done by American bombes. It also agreed to construct only 100 of its own three-wheel bombes, as they had merely half the capacity of the British model. But later on, British four-wheel bombes (needed to break the four-wheel version of Enigma used by the U-boats) had a low serviceability rate and were wholly supplanted, by end-1943, by American bombes.

The concordat with the Navy Department had its price. It led to friction with the War Department [6], which had not been consulted, so that it was not until May 1943, after various missions and high level interventions, that there was a similar, but even more far-reaching, agreement on the Army side. This left all German and Italian military ciphers to GCCS and the Japanese ciphers to the War Department.

As one would expect, there was little progress towards exchanging intelligence with the USSR. Information on how to solve German police ciphers and Abwehr hand ciphers was given to the Russians. Since Ultra revealed that various Soviet ciphers were being read by the Germans, the British intelligence authorities were reluctant to release Ultra to the Soviets even when the source was not disclosed. When they did so, they were yielding to pressure from the prime minister. The history states categorically that there is no truth in the suggestion that the Lucy ring [7] was used by the British to pass intelligence to the USSR. A lot of operational intelligence was sent through the British military mission in Moscow. The Soviets were very reluctant to give anything in exchange, even information about captured German equipment.

One feels that the authors are most at home when writing about naval affairs, including the U-boat war. Indeed two of them (Hinsley and Thomas) worked on naval intelligence at Bletchley. They survey fully the events surrounding the introduction, in February 1942, of the additional wheel into the U-boats' Enigma machines. Even though the development had been foreshadowed in Enigma decrypts as early as the spring of 1941, it was not until December 1942 that GCCS managed to break the four-wheel version (codenamed Shark by the Allies). This notwithstanding that, due to the classic error of an operator using four-wheels prematurely and then repeating the message on three, the wiring of the fourth wheel was recovered in December 1941. Appendix 19 discusses the breaking of Shark and the reasons for the delay, which was to contribute to heavy Allied shipping losses. Those losses were made worse due to the fact that in February 1942 the Beobachtung-Dienst ("B-Dienst", the German Navy's radio intelligence section) finished its solution of Britain's Naval Cypher No. 3, so that they could read most Allied signals about the North Atlantic convoys.

Solving Shark was obviously of considerable importance. The authors do not, however, describe it as the decisive factor in defeating the second U-boat campaign which began against the convoys in December 1942. They do not therefore wholly agree with the German naval historian Jurgen Rohwer, in his assessment that the turning point would otherwise have come "months, maybe many months" after May 1943 [8]. The book also brings out the use of Ultra, in the war at sea, in other ways as well as against U-boats. An auxiliary raider (Komet, or Raider B) was attacked and sunk because its movements were known, largely due to Ultra. Blockade runners were intercepted. Minelaying

operations by Bomber Command were stepped up considerably because Ultra showed them to be especially useful against German coastal shipping. Many other examples are given.

The air war presented a very different picture from the war at sea. U-boat operations, being based on the pack, were controlled by the German submarine commander in chief, Admiral Karl Donitz, through two-way wireless communications. The German Air Force in the West largely used land lines and transmitted little of consequence in Enigma. So there were few nuggets of strategic importance to be gleaned from Luftwaffe Enigma. Sigint and, in particular, the RAF Y service therefore provided much useful intelligence about Luftwaffe operations, including night defenses.

There is full coverage of the war in the Mediterranean and North Africa. The book demonstrates just how important Ultra was in contributing to the defeat of the Axis in North Africa. Important, but the battle of Gazala in May and June, 1942, showed that Army Y intelligence could be more helpful than Ultra during a battle (not least because, being provided locally, it was generally available more quickly) and that it was vital to integrate "Y" fully into the headquarters' structure.

The appendices, totalling about 140 pages, contain some of the most interesting material in Volume II for the general reader, mainly because they are more analytic in tone. Appendix 1 considers British cipher security during the war and gives the reasons for the Admiralty's reluctance to adopt the Typex machine [9], hitherto inexplicable to this reader. Instead it depended on codes (confusingly, a 4-figure code used by officers was called a cypher), which were recoded by a subtractor system. GCCS did not dissent from this decision because it believed that messages would be secure if the recycling tables were replaced at frequent intervals. The volume of traffic prevented that. The real culprit was divided responsibility for the Navy's cipher arrangements, which was split between three different sections.

Appendix 1 also gives an account of the results of German cryptanalytic work on British codes and ciphers. The well-known B-Dienst success against naval codes was considerable. This was especially so for Naval Cypher No. 3, which was used from 1941 to at least the end of 1943 by the United States, British and Canadian navies in the Atlantic. Although there were signs from other decrypts in the second half of 1942 that it had been cracked, it was not until Shark was broken that GCCS realized the full extent of the penetration. There is a short summary of work done against Allied merchant marine and merchant ship codes.

A separate part of Appendix 1 considers the security of Ultra [10]. The reader is left with the impression, from that appendix and the book as a whole, that there may have been too many risks taken, even though Ultra itself showed that the Germans were having doubts about their own cipher security, albeit that the messages did not show their concern, on occasion, about Enigma. More is to appear on this subject in Volume III. How was the arrival of British submarines at Cape Verde during a U-boat rendezvous in September 1941 to have been explained by the Germans? Or the sighting by Allied destroyers of a refuelling between submarines on January 13, 1943? The cover stories must have been very thin. Interestingly, the fourth wheel added to the U-boats' Enigma was a step taken against only "internal insecurity" (the risk of Enigma being read by persons on other Enigma nets [11]).

Appendix 3 is the only place where the names of individual cryptanalysts feature. It sets out in full a letter from Alan Turing, Gordon Welchman, Hugh Alexander and Stuart Milner-Barry [12], all of whom worked in Huts 6 and 8. Following a visit from Winston Churchill to Bletchley, they appealed to him over the heads of officialdom for about 60 to 100 extra staff. Churchill treasured Ultra, which he knew as Boniface. On the same day that he received the letter, he wrote an "Action this Day" minute asking for Bletchley to "have all they want on extreme priority." Their needs were very soon met.

Appendix 4 is the fullest list yet published of various Enigma keys [13]. The naming system adopted by GCCS is given (insects for Fliegerkorps keys, birds for the Army's, fish for the Navy's and so on). By the war's end, about 220 keys had been attacked, of which only 20 or so were naval. The German Army home administration key, used throughout the war, was broken only 13 times and some of those breaks required prisoner-of-war help. A key used by the Gestapo, also for the entire war, was never solved. Although this is described as a classic mystery of Hut 6, both keys illustrate just how secure Enigma could be when used sensibly. Welchman states that it would then have been impregnable [14]. The examples of these keys, taken with the length of time required to crack Shark, may serve as indirect testimonials to the pioneering achievement of the Poles, working as they did with exiguous resources of men, money and machines.

Another appendix covers German police and hand ciphers. A considerable effort was mounted against them, because they provided an entry into Enigma, and much the same system was used by the German Army and Air Force. It also confirms, as first revealed by Peter Calvocoressi [15], that GCCS knew, from spring, 1942, until February, 1943, the precise numbers of deaths in 10 German concentration camps, including Auschwitz and Dachau.

No mistake should be made about the importance of this volume. Quite simply, it is the key to a full understanding of the history of the period. Virtually every other book published so far on the history of World War II in the areas covered by it has to be interpreted in its light. There are no startling revelations, because the general story, including that of Enigma, is now well known. Since most, but not all, of the papers seen by the authors are now in the Public Record Office, there will be other books which may on occasion take a different view to that of the authors or expand the treatment of specific topics. None is ever likely to replace it.

This work will remain a classic and be consulted for very many years to come. For the historian, it is essential reading. The average reader is more likely to want to dip into it from time to time; he will almost certainly profit from doing so. For each, Welchman's book complements this volume on the technical aspects of work on Enigma at Bletchley, as well as filling in some of the human background.

The book is excellently produced and reasonably priced for a specialist work. The proofreading and the index are superb. One looks forward to the next volume (which is to appear in two parts) with a sense of real anticipation, which could not be said after reading Volume I.

No assessment is made in this volume of the ultimate effect of the Allies' ability to read Enigma. Hinsley takes the view that is shortened the war by about three years [16]. Whatever the period, many thousands of lives must have been saved as a result. Ultra could have no better epitaph.

REFERENCES AND NOTES

1. William P. Bundy, "From the Depths to the Heights," Cryptologia, 6 (January 1982), 64-73.

2. House of Commons Debates, 941 (January 12, 1978), cols. 829-830.

3. On the cryptanalysis of Enigma, see Marian Rejewski, "Mathematical Solution of the Enigma Cipher," Cryptologia, 6 (1982): 1-18; the appendix by Tadeusz Lisicki in Jozef Garlinski, Intercept: The Enigma War (London, 1979); Gordon Welchman, The Hut Six Story: Breaking the Enigma Codes (New York, 1982).

   On the breaking of Geheimschreiber by GCCS, see Welchman, 176-179; Brian Johnson, The Secret War (London, 1978), 338-347.

4. For a fascinating and well written autobiographical account of the RAF's Y Service, see Aileen Clayton, The Enemy is Listening (London, 1980).

5. See Alfred J. Price, Instruments of Darkness. (London, 1977).

6. Welchman states that, even as late as 1944, similar difficulties made it impossible for the United States Army and Navy cryptanalytic divisions to discuss Enigma topics with each other directly although they wanted to do so. The British acted as a go-between: Welchman, 74, 75.

7. A Soviet intelligence net in Switzerland that is alleged to have provided quantities of valuable intelligence. Its sources have never been discovered. See further, Garlinski, 107-118; Pierre Accoce and Pierre 'Acquet, The Lucy Ring (London, 1967).

8. Jurgen Rohwer, "Ultra and the Battle of the Atlantic. The German View," Presentation in 1977 to the U.S. Naval Academy, 12. [Cited in Volume II.]

9. On this machine, see Louis Kruh and C. A. Deavours, "The Typex Cryptograph," Cryptologia, 7 (1983), 145.

10. What was perhaps the greatest danger to Ultra's security has received too little attention in this work (only the barest of references in a footnote on page 145 of Volume I). Following the invasion of Poland, the team of Poles who first broke Enigma, including the now famous trio of Rejewski, Zygalski and Rozycki, worked in France under Colonel Gustave Bertrand of the French Army, from October 1939 until November 1942. They were first based near Paris and, after France's defeat in June 1940, in unoccupied France at Uzes (near Avignon). Known as Unit 300, they had to escape when the Germans occupied all of France in early November 1942. Some, including Bertrand and Colonel Gwido Langer of the Polish Army, were captured. At least one, Antoni Palluth, the chief engineer of Unit 300, died in a German concentration camp. It is to the eternal credit of all those taken prisoner that none revealed anything about their work, although some were questioned by the Gestapo: Johnson, 321-323; Christopher Kasparek and Richard Woytak, "In Memoriam Marian Rejewski," Cryptologia, 6 (1982), at 22,23; Garlinski, 132,133.

11. Welchman, 39 and the signal of April 22, M41, from the commander in chief of the German Navy, prohibiting unauthorized tuning in to the U-boat net, in Andrew Hodges, Alan Turing: The Enigma (London, 1983), 199.

12. On each of these, see Welchman, and Ronald Lewin, Ultra Goes to War: The Secret Story. (London, 1978.)

13. On the components of a key, see Welchman, 43, 44.

14. Welchman, 168.

15. Peter Calvocoressi, Top Secret Ultra (London, 1980).

16. The Times [of London], (September 9, 1981), 5.

"Oh, I thought you said it was a cider machine."

FROM THE ARCHIVES

SUBJECT: CODES AND CIPHERS FOR COMBINED AIR-AMPHIBIAN OPERATIONS

[Ed. Note. During the course of their research, our editors and readers are sometimes responsible for the declassification of previously undisclosed material. Or they may discover items in private or public collections, libraries, and archives, items which are not widely known. The purpose of this column is to give these documents wider circulation for the benefit of the cryptologic community. If you have or know about material suitable for this column, please send it to David Kahn, 120 Wooleys Lane, Great Neck, NY 11566. All contributions used will credit the donor.]

[Louis Kruh brings this item to our attention concerning the efforts to coordinate the cipher material for the Allied invasion of France.]

WAR DEPARTMENT The Adjutant General's Office, Washington
AG 311.5 (3-24-43)OB-S-SPSIS-M BJS/reh-2B-939 Pentagon
March 28, 1943
SUBJECT: Codes and Ciphers for Combined Air-Amphibian Operations.

TO: The Commanding Generals,
Army Ground Forces;
Army Air Forces;
Army Service Forces.
Commander-in-Chief, Southwest Pacific Area.
The Commanding Generals,
Theaters of Operations;
Defense Commands;
Departments;
Service Commands;
Base Commands;
Military District of Washington.
The Commanding Officers,
Base Commands.

1. The combined Communications Board, on February 17, 1943, approved the inclosed tabulation of cryptographic systems suitable for use in various stages of combined air-amphibian operations.

2. In the case of each combined air-amphibian operation, it should be the duty of the senior officer appointed for that operation to select the cryptographic systems to be used.

3. The systems set forth in column three of the tabulation should be used if available; if not available, the selection by the senior officer appointed for the operation of any systems in columns one and two is approved.

4. A statement of these cryptographic systems should be included in the Combined Communication Plan for each operation which, after approval by the Commanders, will form the basis for the detailed communication orders of all services concerned.

5. It is desired that necessary steps be taken to disseminate this information in your Command.

By order of the Secretary of War: J.A. ULIO,
Major General, The Adjutant General.

ENCLOSURE A

List of Cryptographic Systems
suitable for
Combined Air-Amphibian Operations

BRITISH	UNITED STATES	COMBINED
1. During Voyage	1. During Voyage	1. During Voyage
Normal naval ciphers* and codes are used for purely Naval traffic and for any essential Army and R.A.F. traffic which has to be passed to the force while at sea. Distribution:-	Normal naval ciphers (ECM* and/or Strip), are used for purely Naval traffic. Joint ECM* and/or Strips for Joint traffic which has to be passed to the force while at sea. Distribution:-	Normal combined ciphers (Combined Cipher Machine C.C.B.P. -and Naval Cipher No. 3) are used for purely Naval traffic and for any essential Army and Air Force traffic which has to be passed to the force while at sea. Distribution:-

BRITISH Normal	UNITED STATES Normal	COMBINED Normal
*Where the CCM is not available, and machines of either service are required to be employed for combined use, liaison groups must accompany this equipment. In some instances where the British Type X machine is selected to be used, the Senior British Signal Officer may authorize its employment by U.S. Forces without a liaison group.	*Where the CCM is not available, and machines of either service are required to be employed for combined use, liaison groups must accompany this equipment. In some instances where the British Type X machine is selected to be used, the Senior British Signal Officer may authorize its employment by U.S. Forces without a liaison group.	
2. <u>During Final Approach and Assault</u>.	2. <u>During Final Approach and Assault</u>.	2. <u>During Final Approach and Assault</u>.
(a) Naval code with a special edition of a subtractor table is set aside as main "high-grade" means of communication at this stage. Distribution:-	(a) Hazardous duty strip ciphers, Navy only, plus Joint Army-Navy Strip Ciphers. Distribution:-	(a) Normal combined cipher as in 1 above except special hazardous duty keys will be employed. Distribution:-
Commander-in-Chief, Flag officers, and Major War Vessels.	Commander-in-Chief, Flag officers, and Major War Vessels.	Commander-in-Chief, Flag officers, and Major War Vessels.
2. (b) For more rapid communication requiring reasonable security for a limited period a 3 letter hatted code is used, unenciphered, known as Combined Operations Code Part I. Distribution:-	(b) For more rapid communication requiring reasonable security for a limited period the Joint Operations Code either unenciphered or enciphered by 4 random mixed alphabets changed at least daily. The Hagelin Machine.	(b) For more rapid communication requiring reasonable security for a limited period, <u>either</u> Combined Operations Code Part I unenciphered <u>or</u> Joint Operations Code, (enciphered by means of

BRITISH	UNITED STATES	COMBINED
	Distribution:-	4 random mixed alphabets changed at least daily), is suitable. Distribution:-
All ships, Main and sub-beach signal stations, Headquarters down to battalions, R.A.F. formation Headquarters and Stations. (point-to-point).	Down to minor war vessels, All Marine activities down to battallions, Army Headquarters down to battalions, Down to Army Air Force Squadrons.	Down to minor war vessels, Main and sub-beach signal stations, Marine Headquarters down to battalions, Army Headquarters down to battalions, Down to Air Force Squadrons.
(c) It is the intention to use plain language considerably in the assault in the interests of speed temporary cover being obtained by the use of code words for important places, names, and references. These code words are included in operational order.	(c) It is the intention to use plain language considerably in the assault in the interests of speed, temporary cover being obtained by the use of code words for important places, names and references. These code words are included in operation orders.	(c) It is the intention to use plain language considerably in the assault in the interests of speed, temporary covering being obtained by the use of code words for important places, names, and references. These code words are included in operational order.
(d) For reconnaissance and air support during the assault stage a non-confidential code is used unenciphered for brevity and standardization, and is known as Combined Operations Code Part II. Distribution:-	(d) For reconnaissance and air support during the assault stage the Joint Operations Code unenciphered and/or the Hagelin Machine is used. Distribution:-	(d) For reconnaissance and air support during the assault stage <u>either</u> Combined Operations Code, Part II <u>or</u> Joint Operations Code unenciphered. Distribution:-

	BRITISH	UNITED STATES	COMBINED
2.	All ships, Main and sub-beach signal stations, Headquarters down to Company F.O.O.'s, R.A.F. Formation Headquarters and Stations Aircraft.	Down to minor war vessels, Army down to battalions headquarters, Naval gunfire liaison parties, Army Air Force down to Squadrons, Aircraft.	Down to minor war vessels, Down to main and sub-beach signal stations, Down to battalion headquarters, Naval gunfire liaison parties, Down to Company F.O.O.'s, R.A.F. Formation Headquarters and Stations, A.A.F. down to Squadrons, Aircraft.
3.	After Establishment of Army Ashore. Normal Joint interservice communications. (a) Type X with interservice settings. Standby: Interservice Cipher (4 figure book with subtractor tables). Distribution:- Navy shore authorities only, Army down to Divisions, R.A.F. down to established R.A.F. Stations.	After Establishment of Army Ashore. Normal Joint interservice communications. (a) Joint ECM. Stand by: Joint Strips. Distribution:- Flag officers and above (afloat and ashore), Army down to Divisions, Army Air Force down to Commands.	After Establishment of Army Ashore. Normal Combined interservice communications. (a) Combined Cipher Machine (C.C.B.P.-). Standby: Combined Strip Cipher System (C.C.B.P.-). Distribution. Flag officers and above (afloat and ashore), Army down to Divisions, R.A.F. down to established R.A.F. Stations, Army Air Force down to Commands.

	BRITISH	UNITED STATES	COMBINED
	(b) Interservice Code (3 figure book with subtractor tables). Distribution:-	(b) Joint Operations Code (enciphered by 4 random mixed alphabets changed at least daily.) Joint Hagelin Machine. Distribution:-	(b) Either Interservice Code with subtractor tables, or Joint Operations Code (enciphered by means of 4 random mixed alphabets changed daily). Distribution:-
3.	Navy down to Destroyers, Corvettes and any other craft which may require to communicate with Army or R.A.F., Army down to Brigades, R.A.F. down to R.A.F. Stations and detached squadrons (for point-to-point use.)	3. Down to minor war vessels, Army down to Battalion, Army Air Force down to Squadrons (for point-to-point use.)	3. Down to minor war vessels, Army down to Battalion (US) or Brigade (Br) Air Force down to Squadrons (for point-to-point use.)
	(c) Air-to-Ground and Air-to-Sea:	(c) Air-to-Ground and Air-to-Sea:	(c) Air-to-Ground and Air-to-Sea: Any of the following may be used:
	Syko or REKOH for interservice use, which is expected to be small, Normal R.A.F. codes for communication with R.A.F. ground stations, e.g. Bomber Code Book.	Hagelin Machine,	Bomber Code Book, Joint Operations Code (unenciphered), Air-Ground Liaison Code, or REKOH for interservice use.

AN UNKNOWN CIPHER DISK

DAVID SHULMAN

The multiple cipher disk illustrated by photographs taken by Louis Kruh is of unknown origin. I purchased it at an antique show about 15 years ago and all I remember the dealer telling me is that he had bought the item with a collection of other instruments in Stockholm from another dealer. Shortly thereafter, I showed it to William F. Friedman in Washington but he did not recall seeing anything of this kind before, nor could he provide any information about it. Perhaps a reader can identify the device and tell us more about who used it and how it was used.

Figure 1.

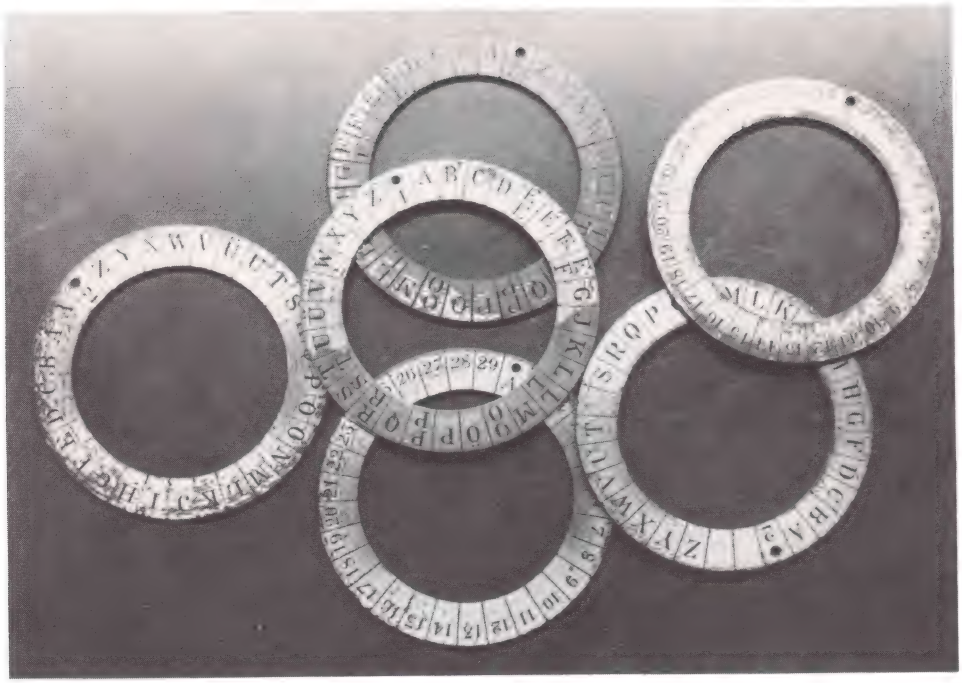


Figure 2.

There are three hinged cases in which the disk and alphabet letters and numbers fit perfectly into the plush lining. Judging by these cases -- I have seen similar ones for holding medals — the device is from the nineteenth century.

Three concentric circles comprise the basic disk which is made of brass. A cursor or metal reading guide is attached to the outer rim. The two outer circles or disks have the numbers 1 - 30 engraved on their base. The inner circle is blank with a raised metal pimple (Figure 1).

There are six metal circular rings (Figure 2), four with letters and two with numbers, each with a hole to fit over the metal pimple, to place in position on the inner disk. Figure 3 shows an alphabet ring in place.

Two cases hold individual letters and numbers made of ivory to place in any sequence in the middle and outer circles. Each of the disks has room for 30 letters or numbers. Figure 4 shows some of them in place.

A knurled brass knob screws into the center of the device and is used to rotate the inner two disks against the outer one.

Figure 3.

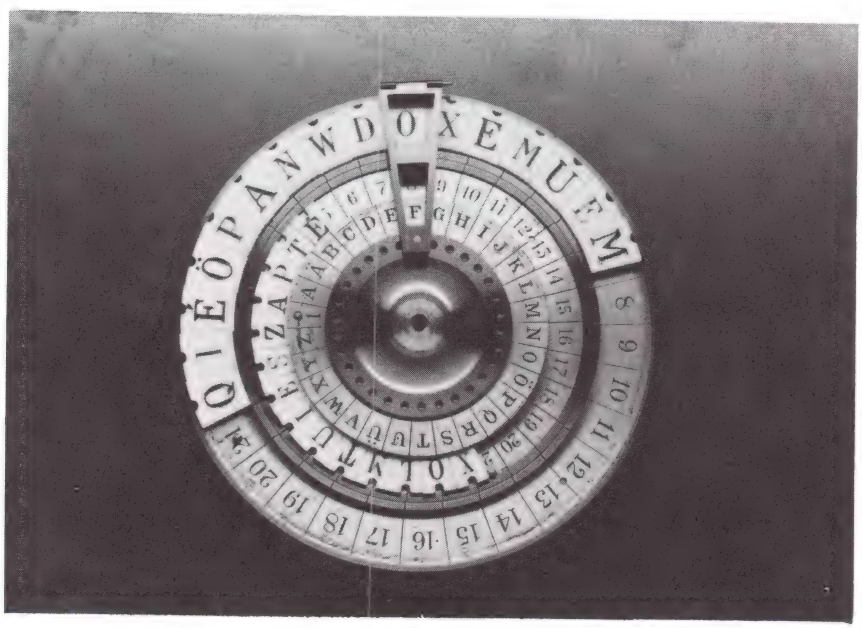


Figure 4.

Louis Kruh asked me to check on what foreign alphabet is represented. My first thought was that it would be Scandinavian, but I found that Swedish, Norwegian, and Danish did not fit, as those languages do not have an alphabet with an E with an acute accent; they also include an O with a slash through it, which this alphabet does not have. Nor could I find any other alphabet of foreign languages to match the disk alphabet. It is possible that these particular letters were provided for someone who might be able to use them in an agreed upon manner between correspondents to represent an alphabet of their own choice, Swedish, Norwegian, and so on.

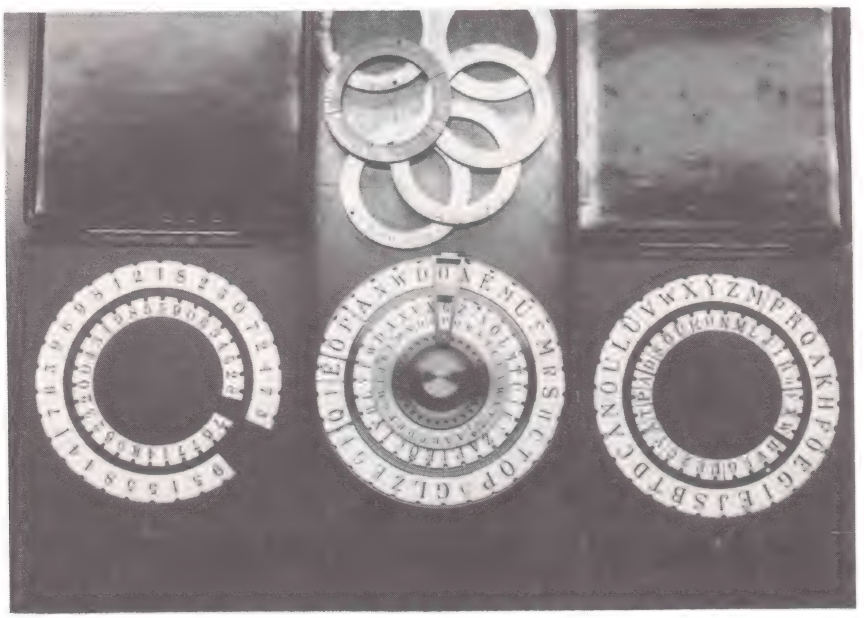


Figure 5. Cipher disk with spare rings and alphabets.

As to the method of encipherment, it could be a keyword polyalphabetic or numerical substitution of any type that can be used with disk encipherments. I believe that the three boxes are complete and were intended for two special correspondents and that they were made to order secretly for high-ranking officials. There is no inscription or mark of manufacture on any of the parts or boxes and no other set is known to exist. I find that the use of ivory in any kind of instrumentation is unusual, same as with precious metals such as gold. It would be, therefore, a good conjecture that what I have was intended for nobility in the Scandinavian countries. It is an elegant cipher disk mystery waiting for someone to shed more light on it, if it is possible.

BIOGRAPHICAL SKETCHES

Charles T. Retter is a principal engineer in Central Processor Development at the Data General Corporation. He received his BS from Drexel University, MS from Northeastern University, and PhD from Johns Hopkins University. He was the principal designer of the Nova 4/C and Eclipse S/280 computers. His research interests are primarily in the area of error-correcting codes. Address: Data General Corporation, B135, 4400 Computer Drive, Westboro MA 01581.

Louis Kruh is a public relations executive with the Bell System in New York City. His interests in cryptology span more than forty years. He collects crypto material and machines. He has a BBA, cum laude, from the City College of New York, and an MBA, with distinction from Pace University. Currently he is nearing completion of law school. Address: 17 Alfred Road West, Merrick NY 11566.

Donald Davies is a scientist at the UK National Physical Laboratory working on matters of data security and authentication. He began to work with digital computers in 1947, helping to build an early machine and then using it for traffic simulation and other studies. By 1965 he had moved to computer networks and developed an early packet switched network (he coined the word "packet"). Eventually this work led to a "distinguished fellowship" of the British Computer Society. Now his professional interests include the DES, public key systems and protocols for their use and his private interests include historic cipher machines. Address: Division of Information Technology and Computing, National Physical Laboratory, Teddington Middlesex TW11 OLW, England.

Borge Tilt is a consultant in statistics and operations research. A native of Denmark, he received the MA degree from the University of Copenhagen in 1959. He lived in the United States during the period 1965-76. In 1976 he received the PhD from the Georgia Institute of Technology in queueing theory. his articles have appeared in Applied Probability, European Journal of Operations Research and other journals. His interest in statistical aspects of cryptology dates back to the reading of David Kahn's The Code breakers. Address: Carl Jacobsens Vej 4 D, 2500 Valby, Denmark.

Ralph Erskine is a lawyer in Northern Ireland. A member of Gray's Inn, London, he graduated from Queen's University, Belfast, in 1955. His interests related to cryptology include microcomputers in general and the British Broadcasting Corporation's Model B, in particular. Address: 25 Hawthornden Road, Belfast BT4 3JU, Northern Ireland.

John M. Carroll is a Professor in the Computer Science Department of the University of Western Ontario. In 1982-83 he was a computer scientist in the military message automation section of the Naval Research Laboratory, Washington DC. During 1975-76, he was consultant to the Electronic Data Processing Security section of the Royal Canadian Mounted Police, where he prepared draft standards on computer security for the federal government of Canada. During 1970-72 he was consultant to the federal Privacy and Computers Task Force. His 1969-70 study on the privacy and security of university student records, prepared under the patronage of the Canadian Council, was one of the earliest efforts in the field. Dr. Carroll is the author of Computer Security, Confidential Information Sources: Public and Private, and Controlling White Collar Crime. His research interests include automatic analysis of documents, computer-aided risk analysis of computer systems, computer simulation of municipal protective services, and data protection for microcomputers. Address: Computer Science Department, University of Western Ontario, London Ontario, N6A 5B9 Canada.

Pierre G. Laurin did graduate work, under the direction of Dr. John Carroll, in Computer Science. Address: Computer Science Department, University of Western Ontario, London Ontario, N6A 5B9 Canada.

David Shulman is a careful researcher in the field of cryptography and possess one of the finest collections of cryptographic material. His research and care for detail over the years has produced his book, An Annotated Bibliography of Cryptography. Long a contributor to the American Cryptogram Association, the National Puzzler's League and The Oxford English Dictionary Mr. Shulman continues to find new material in cryptographic records and bring them out for crypto enthusiasts. Address: 34 Park Row, New York NY 10038.

SUBSCRIPTION INFORMATION

CRYPTOLOGIA is a quarterly journal with issue dates of January, April, July and October. The four journals issued each year constitute one volume. The January 1983 issue is Volume 7, Number 1.

Subscription prices (U.S. Dollars): $28.00 per year for U.S., $36.00 per year for non-U.S. Air Mail overseas rate is $60.00 per year. A subscription begins with the current issue as of date of receipt of request unless otherwise instructed. Back issues from January 1979, Volume 3, Number 1 to current issue are available from the Editorial Offices for $8.00 each in the U.S. and $10.00 each to non-U.S. address. Specify volume, number and issue date.

All orders, checks and inquiries should be sent to: CRYPTOLOGIA, Rose-Hulman Institute of Technology, Terre Haute, Indiana 47803, USA. Make checks payable to CRYPTOLOGIA.

Note to subscribers: The number in the upper right corner of your address label indicates the last issue of your subscription. The right hand (single) digit indicates the Number and the remaining left hand digit(s) indicate the Volume of the last issue in your subscription. Renew your subscription now.

CALL FOR PAPERS

CRYPTOLOGIA welcomes articles on all aspects of cryptology. We especially seek articles concerning mathematics and computer related aspects of cryptology. Articles describing new cryptosystems and methods of cryptanalysis of cryptosystems, historical articles, memoirs and translations are all sought.

Send mathematical and computer related papers to Brian J. Winkel, Division of Mathematics, Rose-Hulman Institute of Technology, Terre Haute, Indiana 47803.

Send papers, inquiries and letters concerning cryptographic machines, devices and equipment to Louis Kruh, 17 Alfred Road West, Merrick, NY 11566.

Send historical and other nontechnical articles to David Kahn, 120 Wooleys Lane, Great Neck, NY 11023.

Any paper may also be sent to the Editorial Office, CRYPTOLOGIA, Rose-Hulman Institute of Technology, Terre Haute, Indiana 47803.

Three copies should be submitted and one should be kept by the author as a protection against loss. Manuscripts should be legibly typewritten, or reproduced from typewritten copy and double-spaced with wide margins. All papers should have an Abstract and a Key-Word List after the title and author. Editorial style follows the University of Chicago Press Manual of Style. Please adhere to the footnoting style found in CRYPTOLOGIA articles. Diagrams should be done in black, suitable for off-set photo reproduction, and clearly labeled with a legend. Photographs should be clear and glossy. Indicate whether or not the photo print enclosed is to be returned.

While the ultimate responsibility for the accuracy of the material presented lies with the author(s), the Editorial Office will do its best through the refereeing and consultation process, to help insure correctness.

Authors will receive two copies of the issue in which their articles appear.

CRYPTOLOGIA

A Quarterly Journal Devoted to All Aspects of Cryptology

Volume 8 Number 2 April 1984

Table of Contents

Published Quarterly at

Rose-Hulman Institute of Technology

Terre Haute, Indiana 47803 USA